McGRAW-HILL SERIES IN PSYCHOLOGY

CLIFFORD T. MORGAN, *Consulting Editor*

Personnel and Industrial Psychology

McGraw-Hill Series in Psychology

Clifford T. Morgan, *Consulting Editor*

Barker, Kounin, and Wright · Child Behavior and Development
Bartley · Beginning Experimental Psychology
Blum · Psychoanalytic Theories of Personality
Brown · The Psychodynamics of Abnormal Behavior
Brown and Ghiselli · Scientific Method in Psychology
Cattell · Personality
Crafts, Schneirla, Robinson, and Gilbert · Recent Experiments in Psychology
Deese · The Psychology of Learning
Dollard and Miller · Personality and Psychotherapy
Dorcus and Jones · Handbook of Employee Selection
Ferguson · Personality Measurement
Ghiselli and Brown · Personnel and Industrial Psychology
Gray · Psychology Applied to Human Affairs
Gray · Psychology in Industry
Guilford · Fundamental Statistics in Psychology and Education
Guilford · Psychometric Methods
Haire · Psychology in Management
Hirsh · The Measurement of Hearing
Hurlock · Adolescent Development
Hurlock · Child Development
Hurlock · Developmental Psychology
Johnson · Essentials of Psychology
Karn and Gilmer · Readings in Industrial and Business Psychology
Krech and Crutchfield · Theory and Problems of Social Psychology
Lewin · A Dynamic Theory of Personality
Lewin · Principles of Topological Psychology
Maier and Schneirla · Principles of Animal Psychology
Miller · Experiments in Social Process
Miller · Language and Communication
Misiak and Staudt · Catholics in Psychology: A Historical Survey
Moore · Psychology for Business and Industry
Morgan and Stellar · Physiological Psychology
Page · Abnormal Psychology
Reymert · Feelings and Emotions
Richards · Modern Clinical Psychology
Seashore · Psychology of Music
Seward · Sex and the Social Order
Shaffer and Lazarus · Fundamental Concepts in Clinical Psychology
Siegel · Nonparametric Statistics: For the Behavioral Sciences
Stagner · Psychology of Personality
Townsend · Introduction to Experimental Method
Vinacke · The Psychology of Thinking
Wallen · Clinical Psychology: The Study of Persons
Zubek and Solberg · Human Development

John F. Dashiell was Consulting Editor of this series from its inception in 1931 until January 1, 1950.

Personnel and Industrial Psychology

Edwin E. Ghiselli

PROFESSOR, DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA

Clarence W. Brown

PROFESSOR, DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA

SECOND EDITION

McGRAW-HILL BOOK COMPANY, INC.
New York Toronto London
1955

PERSONNEL AND INDUSTRIAL PSYCHOLOGY

Library of Congress Catalog Card Number: 54-10632

Preface

It is unfortunately true that advancements in the development and use of scientific personnel and industrial methods are greatest under stressful social conditions. At these times attention is focused upon scientific methodology and principles in order to achieve solutions to urgent problems in manpower utilization and production. After the immediate emergency has passed, whether it is an economic depression or war, some of the scientifically derived procedures are continued in force, with the result that business, industry, and labor employ to a greater extent sounder procedures in dealing with human problems.

While the rate at which scientific methods are adopted is, in the minds of many industrial psychologists, so slow as to cause despair, it is nevertheless true that if a long-term view is taken significant improvements will be revealed. It is difficult to measure exactly and in quantitative terms the trends toward the use of scientifically developed principles and procedures. Yet it is apparent that there has been increasing application of such procedures and principles, *e.g.*, the increased use of psychological tests in the selection and placement of workers, and the greater attention given to employee training and to problems in human relations. Certainly if our universities have had any effect whatsoever upon their students, the thousands who have been exposed to courses in industrial psychology and personnel methods must have carried away with them some understanding of the usefulness of scientific methods and findings, which later contributed to their thinking about the personnel problems with which they were confronted.

In writing this book the authors have, of course, emphasized the need for using scientific methods and procedures. The choice of topics discussed no doubt is biased by the kinds of specific experiences that they have had in working with personnel problems in business, industry, and government. In this second edition attempts have been made to give a better-balanced view of the field, particularly with respect to the areas of motivational and social factors. The reader will find that certain approaches and emphases have broken with the traditional, and attempts have been made to reorient the thinking and analysis in terms of what appear to the authors to be the primary problems of the future.

It is the authors' hope that this book will meet the needs both of university students and of those in business, industry, and government who desire to obtain a foundation in the basic principles involved in the application of scientific psychology to manpower problems. The extent to which this hope will be realized will, in part at least, be a function of the degree to which the reader will be willing in the beginning to consider basic underlying principles and to adopt a critical attitude before seeking solutions to his immediate problems.

It is difficult to acknowledge in any satisfactory manner the aid the authors have received from others in the preparation of this book. Appreciation is expressed for the assistance rendered by the many students at the University of California who critically evaluated the first edition of the book and its four earlier mimeographed editions. Further recognition must be given to students and to management and labor personnel who critically evaluated many of the new ideas and approaches contained in this second edition. We wish specifically to thank our colleagues Mason J. Haire and Ralph R. Canter, and Col. A. C. Tucker, USA, who kindly reviewed many of the chapters of this second edition. Finally, the authors wish to acknowledge the generous assistance rendered by their wives, to whom this book is gratefully dedicated.

EDWIN E. GHISELLI
CLARENCE W. BROWN

Contents

CHAPTER 1

The Scientific Study of Human Factors in Industrial Problems

There are few developments in history that match in scope and importance the growth and expansion of industry. Every year brings new discoveries which enable man to extend further his control over nature, thus satisfying a greater proportion of his wants. The most rapid development has been in the invention and exploitation of changes in the physical processes of production. Some progress has also been made in discovering psychological factors which contribute to effective production. In this area, however, many problems remain to be solved before the most appropriate psychological climate for work can be established.

GROWTH OF INDUSTRY TOWARD A DEMOCRATIC ORGANIZATION

In the development of industry in the free countries of the world the growth has been in the direction of what for want of a better term can be called a democratic industrial organization. By this is simply meant that the directions taken by industrial expansion have been determined to a large extent in terms of the needs of all the people and, further, that in greater and greater degrees more and more people have come to share in a personal way in the responsibility for the progress that is being made.

Growth of the Physical Phases of Production. Physical scientists, working in university and industrial research laboratories, have been primarily responsible for the nature and rapidity of industrial development. The so-called pure or theoretical scientists have led the way, moving ahead with little thought of the practical benefits that might accrue from their findings. Closely following have come the industrial scientists and engineers who, with their understanding of both science and industry, have carried the theories about nature into the practical world and reorganized the physical environment to serve human needs.

Industrial progress has advanced along two general fronts. Advance-

ments have been achieved in a variety of areas where control of the physical environment has been made to serve man's needs. Further, many technical procedures have been developed by which the cost of production has been greatly reduced and the amount of production greatly increased, making it possible for industrial advancements to be enjoyed by an ever-increasing proportion of the population. So-called modern conveniences are to be found in every area of human living. In the home, at the factory workbench, on the playground, in the church—actually in every activity of man—are found practical devices of many descriptions which bring the physical environment under his control. And these conveniences are no longer available only to the few. Laborsaving machinery, production-line procedures, and a closely knit organization of the channels of distribution from the raw natural elements to the finished product in the salesman's showcase have accomplished the multiplication of manufactured items to a point where everyone shares in some of the advances and most people share in a large way.

Development of Emphases on Human Factors. One of the most prominent and persistent facts that has emerged from the rapid development of industry is the importance of the role played by the worker. Some of the earlier scientists did not restrict their interests to the physical conditions of work but became interested in the worker himself. They saw the necessity of studying the human problems of production and endeavored to apply scientific methods to their solution. Many sound solutions were effected by them. The over-all success attained by these earlier investigators of human factors, however, did not rival the achievements of the physical scientists.

In more recent times the human factors in industry have become the center of an increased amount of research. Despite the tremendous strides made in the discovery of laborsaving machinery there has been no lessening of the need for trained personnel. The unionization of workers and the scarcity of qualified personnel have been the two great forces which have brought about awareness of the significance of labor's contribution to industry and the acknowledgment of the nature and amount of the rewards that should be returned to labor for this contribution.

One of the most remarkable advances in the human phases of industrial production has been the recognition of the complex relationships that exist between the employee's work and the many other phases of his living. The worker is no longer treated as a mere cog in the production machinery. His personal and social activities touch upon his job performance in many intricate ways. In turn, his job touches upon his personal and social life in many respects. Self-satisfaction from achievement, individual feelings of pleasure from creative work, and fulfillment, through the job, of immediate and remote personal and social goals are part of

labor's stake in industry. Management is becoming increasingly aware of the significance of the psychological values which are to be obtained from the job, and the ways in which these values reach out into all aspects of the worker's life.

The reason for incorporating these human phases in the job picture is twofold. Such factors are conducive to an increased contribution by labor to the end products for which industry is organized. But more important, the adjustment of working conditions in recognition of these human factors is coming to be accepted as a bona fide part of the return which industry itself is obliged to make to the worker. There is, then, a standing and continuous need to adjust the work to the needs and abilities of the worker and to assist the worker to adjust himself to the needs and demands of his job. This new interrelationship is replacing the older policy under which the work remained constant and the worker was forced to make all the adjustments if he were to survive on the job.

Social Effects of Industrial Development. The expansion of industry has had profound effects upon social conditions. For present purposes only three general types of effects will be considered, namely, changes in the status of the worker, changes in the nature of the organization of business, and changes in the general welfare of society.

The effects of industrial growth on the worker's status were alluded to in previous discussions. Years ago the worker was mere chattel and was dealt with much like a piece of machinery on the production line. His only function was to produce and the more he produced the better worker he was considered to be. The objective was to effect the greatest output with the least cost, and many techniques and procedures were introduced with the sole purpose of making the human machine turn out more work.

Happily this point of view is outmoded in democratic societies. Slowly, oftentimes painfully, and at great cost, the employee's status has been changed. Most workers are no longer required to labor 12 or 14 hr. a day. In an increasing number of industrial organizations the worker cannot be fired or transferred except for due cause. The chances of being injured on the job or of incurring some other type of occupational disability have been greatly reduced. If a worker is injured, there is a fair chance that he will receive some compensation for his disability. Through workers' welfare funds, pensions, and old-age security benefits, more workers and their families are gaining economic security against the day when the breadwinner will no longer be capable of work. These benefits are illustrative of some of the gains derived from championing the worker's part in production.

During the growth of industry many changes have occurred in the nature of the organization of industrial concerns. Some of these changes

have been made within companies, and some between companies. Both kinds have produced effects of social significance. The assembly line can be mentioned as an example of the first kind of change. Here through an ordered arrangement of several jobs associated with the assembling of a common item the rapidity of manufacture is greatly increased and the cost per item greatly reduced. But along with these beneficial effects certain detrimental effects also occur. Sometimes the task assigned to a given worker is so simple and repetitive that it becomes extremely monotonous. Sometimes speed-up methods set the speed of the assembly line at a level which is close to the tolerance limits of some of the workers.

As illustrative of the organization between firms, the familiar holding companies may be mentioned. Like the assembly line, the effects of holding companies are "mixed blessings." Holding companies often create monopolies and eliminate free competitive markets. Thus artificial limits are set on the amount of production, and prices to the consumer are arbitrarily established. On the other hand, it is through a tightly knit organization of the different phases of manufacture from raw material to finished article that costs of production are reduced to the point where the wage earner can afford to enjoy the manufactured article. Furthermore, it is through the increased breadth of facilities found in large companies that some of the most important research problems facing society can be tackled and a free enterprise system still retained.

The effects of industrial expansion on the general welfare need not be belabored. All the changes described above have had repercussions beyond the effects upon the immediate employers and employees concerned. The continued rise in the standard of living is testimony to the profound changes made in the general welfare. Today in industrial democracies is found the highest standard of living yet attained by man. As mentioned above, not all the effects upon the general welfare have been beneficial. But the occurrence of detrimental effects simply means that the task is not finished. They raise a challenge for the future and demand that research in the human factors be accelerated so such effects may the sooner be eliminated.

Joint Responsibility of Labor and Management in the Future Development of Industry. An important lesson from the past concerns the need for joint responsibility between labor and management for the maintenance of a sound economy. Industry does not exist solely for either the employer or the employee. It must furnish adequate incentives for both. The task of the employer is one of discovery, organization, and long-term planning. The task of the worker is to manipulate the materials, run the machinery, and operate the processes that manufacture the goods. One complements the other. Only by supplying adequate incentives for each can their cooperation in the joint enterprise succeed.

An evaluation of the human factors in industrial production is necessary for this cooperation. The understanding of these factors has lagged behind the understanding of the physical factors of work. An important emphasis of the future is the same objective and thorough application of the scientific method to the human side of work which has been applied to the physical side. This will be a cooperative enterprise of scientists from several fields. Certainly the industrial psychologist will be expected to make a significant contribution.

THE NATURE OF THE SCIENTIFIC METHOD

No one will dispute the fact that industrial progress has come about, either directly or indirectly, through the contributions of science. Growth has resulted from successive technological improvements that have issued primarily from the research activities of physicists, chemists, and engineers. Scientific method has increasingly furnished the framework on which industrial production has been built, and its concepts have permeated nearly all phases and branches of industry. If more rapid progress is to be made in the solution of the human problems in industry it will come through a more intensive application of the principles of science. It might be well at this point to review briefly the characteristics of the scientific method.

Science and Common Sense. The reader unfamiliar with the scientific method will not find its procedures entirely different from those he uses in the method of common sense. Many features of the two methods are alike, which leads the uncritical to confuse them. Some of the major differences are matters of degree, of extent, and of adequacy, rather than matters of kind. It is in terms of such differences, however, that science has achieved tremendous progress and that, in comparison, common sense has attained only moderate gains. In the following descriptions of the characteristics of science the reader will be able to note those features in which the two methods vary significantly.

Facts Are the Basis of Science. The scientist predicates all his judgments and evaluations on fact. A fact is defined broadly as being any occurrence, happening, experience, or phenomenon. One of its most important characteristics is that it is impersonal. This means that a fact, although dependent upon human observation, is independent of any particular observer. A fact can safely be used without knowing its discoverer, as contrasted with an opinion, which can be used wisely only when the qualifications of the person expressing it are known. Facts are the backbone of the scientific method.

Science Emphasizes Objective Observation. Science endeavors to study its phenomena objectively and only appeals to subjective observation

when compelled to do so. With objective observation, more than one person can experience a given phenomenon. There can then be a duplication of the experience, and one observer can check upon another. Certainly several observations are better than one, especially when the situation is complex and not all its features are fully understood. Agreement among observers is an important condition of the scientific approach and it is more readily obtained with objective than with subjective observation.

Science Is a Systematic Rather than a Casual Study. The study of a problem by scientific procedures requires a systematic attack. The problem is comprehensively investigated from as many angles as possible. One of the important goals of the scientist is to get all the facts. Therefore he will leave no stone unturned in his effort to obtain a complete description of the problem. It is analyzed into all its parts, and the ramifications of these parts are traced and studied. The influence of each factor upon the others is determined, and the parts are assembled and patterned again into the whole. No aspect of the problem is neglected, and each assumes its rightful importance in the composition of the completed picture.

Science Utilizes Experimental Controls. To gain knowledge concerning the influence of a given factor, the scientist utilizes various forms of control that enable him to check on the operation of the factor when it is freed from the possible effects of other related factors. Frequently the success of an investigation is directly determined by the degree of control attained over the relevant variables, and often more effort is expended in devising means to attain this control than is spent in actually conducting the study. The scientist takes advantage of the experimental method to set up studies on various phases of his problem which are conducted under the conditions prevailing in the practical situation, in so far as this is possible. As contrasted with the practical situation, the experimental situation provides more adequate control over various potentially disturbing factors which might otherwise operate to influence the results.

Science Utilizes Quantitative Description. Measurement is one of the most important features of the scientific approach but a difficult one to achieve in the field of human behavior. It forms the basis of prediction, and without it there is little left on which to determine the expectations of the future. Quantitative scores or measures contain the answers to the scientist's problems, but usually they do not yield these answers without considerable manipulation. The quantitative measures by themselves have few significant meanings. Their full import is gained only when they are analyzed and synthesized into arrangements of various complexity. This process of evaluating the quantitative data involves the utilization of many techniques of a statistical nature.

Science Demands Trained Personnel. The method of science implies a scientist to conduct the operations. The full meaning of this statement is not always realized. The scientific approach does not consist of a few simple rules that, when followed meticulously, will produce the needed facts. Scientific method is as sensitive and precise, and as reliable and comprehensive, as the training, ability, and dependability of the scientist. Its correct application depends upon trained personnel, and this requirement is as true in the field of human behavior as it is in the field of physical science.

Science Is a Self-corrective Method. In contrast to other approaches the scientific method provides for correction of its results. Generalizations must be based upon facts, and at all times an attitude of doubt and questioning is encouraged. No claim to infallibility is made. Any new evidence whether positive or negative is incorporated into the picture, and modifications made if the new facts demand them. Explanations are finally accepted because there are no better alternatives. He who applies the scientific method hesitates to rest his case at any time but continuously endeavors to improve the findings through seeking further evidence.

SCIENTIFIC PSYCHOLOGY IN INDUSTRIAL PROGRESS

A description of some of the contributions that scientific psychology is making to industrial progress will form the main portion of the chapters to follow. At the outset, however, it will prove of value to consider briefly the form in which these contributions are being made.

Psychology Offers Few Ready-made Solutions. Scientific psychology has few specific ready-made solutions for the human problems of industry. There are two primary reasons. First, the complexity of the causal structure underlying human behavior is so great that accurate generalization is difficult. Human variability must always be accepted as a bête noire whenever a solution is generalized from one situation to another. The behavior of the employer or the employee is not determined by a few simple factors. On the contrary, it arises from an intricate interplay of a large number of complex factors. A further complication is added in that these factors are not constant in time, but vary with the changing conditions of the individual's responses and the work situation. It is then not to be expected that a given rule of behavior discovered by the psychologist can necessarily be applied with assurance of success to the behavior problems arising on the job.

Secondly, accurate generalization is made difficult because conditions found in the factory, office, or salesroom in which the generalization is expected to apply differ radically from the conditions of the psychologist's laboratory. For example, suppose a new work method is studied in the

experimental laboratory and found very effective. There is no assurance that the method will meet with similar success in a factory. The determinant factors are not identical in the two situations and actually may be very different. Certainly the intangible motivating factors activating the individual in the experimental laboratory are not the same as the motivating factors activating the worker in the factory. There is a ready solution to this difficulty, and that is for the scientist to perform his research under conditions duplicating those of industrial practice. His findings then would have direct application. One way of accomplishing this is to set up research groups within industry as has been done by several organizations.

Psychology Offers Many Hypotheses. The best substitute for a ready-made solution is a hypothetical one which is supported by evidence indicating that it has a high chance for success. For any given hypothesis the chances for success are dependent upon many factors, the most important of which are the following:

1. The degree of similarity between the practical situation and the laboratory situation in which the hypothetical solution was evolved
2. The accuracy with which the laboratory studies were executed
3. The amount of evidence supporting the solution that can be obtained from related studies and experiments

Scientific psychology contains hundreds of thousands of experiments covering a wide range of situations involving human behavior. Nearly every conceivable human factor has come under study at one time or another, and for some factors there are large numbers of experimental findings obtained in a wide variety of different situations. This storehouse of material is a rich source of ideas that will prove useful when applied to the human problems of industry.

It must be borne in mind, however, that any generalization that is selected for trial must be accepted as only tentatively true. The adaptation of any finding to a specific industrial situation is made with a skeptical eye, and careful checks are made of its success under the new conditions. Only when evidence is gained that the finding is successfully operating under the specific practical conditions can it be advanced as deserving of wider adoption in industry.

Psychology Offers the Scientific Method as Adapted to the Study of Human Behavior. The application of the scientific method to any subject matter requires various adaptations and modifications which are imposed by the nature of the subject matter itself. The broad framework of science remains the same, and the larger steps of the method continue practically undisturbed, but the individual processes and the specific techniques are adjusted to the subject. They take on the form most suitable to the particular situations or conditions to be studied.

Scientific psychology can offer the method of science as it has been specifically adapted to the study of human behavior. Within the framework and general outline of the method many specific procedures and techniques have been devised in attempts to subject a variety of behavior determinants to scientific scrutiny. These special adaptations have to do with every phase of the method. Some are involved in setting up the problem to be studied, others in isolating the variables to be measured. Some are involved in the introduction of controls for separating the effects of related factors. Still others are involved in analyzing the data and statistically testing the significance of hypotheses. Most of these specialized procedures and techniques will prove of value when applied to industrial problems.

Psychology Offers Trained Personnel. The scientific method can be accurately used only by persons especially trained in its application. Many mistakes of the past can be traced to the inexpertness of those who conducted the studies. Most industrialists agree that the engineering phases of production should be in the hands of trained engineers. Further, they believe that, as these phases become more complex, the qualifications which the individual must have in order to hold the job should become more technical. Unhappily, this has not been the rule usually followed in regard to the human phases of industry.

For the most part, trained personnel must come from the colleges and universities. In past years academically trained psychologists have not entered industry in large numbers. This condition is changing, and it is expected that, in the future, the challenge of industrial problems will attract to the field larger numbers of these highly trained scientists and technicians. Furthermore, special courses of study are being instituted in some of the larger schools to give more proficient training in industrial problems. It is to be remembered, however, that upon graduation from college this trained personnel will not be fully prepared immediately to conduct research programs. No one knows better than the scientist himself that the successful solution of any problem involves not simply knowledge of a method but also a firsthand acquaintance with the factors conditioning the special problem to be solved. Academically trained personnel must supplement their classroom preparation with firsthand experience with jobs and processes in industry. They must have time to acquaint themselves with the factors functioning in the practical situation. This points to the need for internship, since only after the scientist has gained this practical experience can he be expected to apply in expert manner his special knowledge of psychological techniques to the complex human problems of production.

Psychology Offers Impartial Research. Scientific research means unbiased research. It is difficult to see how a scientist can long be unbiased

when he knows that his job depends upon getting results that favor one of two parties, especially when the parties may have somewhat conflicting objectives. This points to the need for joint sponsorship by management and labor of research on the human problems of industry. Except for a few instances, this is only a hope at the present time, but it should become a widespread reality in the not too distant future.

Impartiality in research is not likely to be maintained when the problems to be investigated are determined by either the employer, on the one hand, or the employee on the other. It is true that, as a method, science can be used to investigate any problem. But it is also true that the selection of the problems to be investigated will determine in great measure the particular applications that are formulated from the results of the research. Regardless of the particular individuals who will be affected, in solving a problem the scientist studying human factors should be free to investigate any and all relevant problems so that all facts will be available for the formulation of any solution.

As a price for its answers, science demands unbiased research. There are many points throughout a research project at which the subjective judgment of the scientist enters as an important contributing factor. A scientist's judgment may be biased when his point of view is restricted by some allegiance. As the aim of the scientific psychologist in industry is to solve the human problems, he can best do this when he is free to follow the lead of facts and not of bias.

Psychology can provide scientists trained in objectivity and impartiality. In the atmosphere of the university the scientist has but one allegiance and that is allegiance to the truth. If he can be said to be biased, he is biased in favor of objectively studying all facets of a problem, of postponing his generalizations until all the facts are in, and of allowing the facts, and not his biases, to determine the nature of these generalizations. With his type of training and background the scientific psychologist can perform impartial research when he enters upon a study of the human factors in industry.

SOME CORE PROBLEMS IN SCIENTIFIC DESCRIPTION

Description occupies a central position in the method of science. An understanding of the factors affecting industrial production, whether these are physical or psychological, is achieved only after the observed events are accurately described and intensively analyzed and studied. Description makes possible a permanent record of the events, thus furnishing a way by which these events can be later manipulated in order to evolve a solution to the problem at hand. If the outcome of any analysis is to be sound it is necessary for the scientist to begin with accurate descriptions.

Need for Quantitative Description. Numerical or quantitative description is of prime importance in scientific psychology. Accurate identification of the events being described, and accurate estimation of the variation and interrelationships of these events are best accomplished when the variables are expressed quantitatively. Greater precision of meaning is then obtained, and disagreement in meaning among different observers is minimized.

Every phase of industry touching upon human performance, whether it be the performance of the employer or of the employee, may at some time need to be accurately described. For example, because the design of a tool affects the productive performance of the worker, tools of different design must be evaluated. Oftentimes in a given task, any one of several makes or designs of tools can be used. To determine which one enables the worker to do his job with the least effort or in the shortest time requires that the several tools be studied in the practical situation and their relative effectiveness determined quantitatively. Similarly, the method of work will condition the effectiveness of the worker. There are usually several ways of doing a job. Some of these may be definitely inferior to others. It is necessary to discover which method will result in maximum performance and at the same time be adaptable to the abilities and experiences of the individual worker. This is done best when the end products of the work methods are compared on a quantitative basis.

Evaluations must be made of the human factors in production. The aptitudes, abilities, and interests of the worker are other areas in which quantitative descriptions are important. It is necessary that each worker be placed on a job where he can work effectively in terms of his qualifications. When these qualifications are expressed in quantitative units, more accurate descriptions are possible, differences between workers are brought into sharper focus, and a more accurate placement of each worker is accomplished.

Description of Average Performance. It is not to be expected that a given factor under study will remain constant under all conditions of operation. It is necessary then to represent the several values obtained for a factor in terms of an average. The familiar arithmetic mean suffices for most purposes.

The evaluation of a factor, whether it is a physical aspect of the working environment or a proficiency characteristic of the individual worker, is accomplished by making comparisons. The importance of any physical factor is determined by comparing the results obtained when the factor is operating with the results obtained when it is not operating. The importance of any proficiency factor is determined from the relation it bears to the productivity of the worker, and this requires the comparison of the productivity of workers who differ in terms of the proficiency factor.

The question then arises: When a factor has more than one value, which of these values should be used in the comparisons? Certainly it is desirable that the most representative value be used so that accurate comparisons will result. The arithmetic average is the most representative figure of the group of numbers from which it is computed and so forms an accurate basis for comparisons.

Measures of average performance are needed for all factors which touch upon job behavior. This means that tools, working conditions, methods of work, length of working spell, accidents, and other similar factors must be evaluated in terms of averages. Of course, it also means that many aspects of the worker's behavior must be described quantitatively and represented in the form of average performance. Evaluations of both physical and psychological factors operating on the job make it possible to detect and isolate those particular factors which are contributing to some given problem.

Description of Variability in Performance. The variability in performance mentioned in the previous topic may be an important factor in its own right. The amount of variation in a machine, in a piece of equipment, in a work process, or in a worker's performance is not a constant value over time. The variability of a given machine is a function of the amount of its use, the quality of its constituent materials, the precision of its manufacture, and the skill and care of the operator. The amount of variability in a worker's performance is a function of his knowledge and skill, his interest and enthusiasm, his goals and ambitions, his satisfaction with his home life and social outlets, and many other similar factors.

The problem frequently occurs of comparing the variability in performance of different machines or of different individual workers. More than one copy of a machine will frequently be in use, and sometimes differently designed machines will be operating side by side. Similarly, the same job tasks may be executed by many workers. It is not to be expected that different machines even if of the same manufacture will give the same quality or quantity of performance. For example, in a printing establishment one might hear the statement, "Sure this press can turn out the work but you never know when it is going to clog up; the other slower machine is more dependable." Similarly, different workers plying the same trade do not produce equal or constant qualities or amounts of work. Some workers are said to be "hot weather" workers, meaning that their production is much better in the warmer months of the year. Some are called "Wednesday merchants" because they seem to get out more work in the middle of the week. Variation in performance is an important characteristic to assess in the evaluation of machines, workers, processes, techniques, materials, etc.

Measures describing machine or human variability in performance may

reveal sources of trouble. For example, wide swings in the amount or quality of performance are symptomatic of a potential malfunction. Wide variations in output on a punch press may be reflective of some adverse condition which might grow worse if not soon corrected. The trouble may be in the machine, in the operator, or in the combination of the particular machine and the particular operator. Through an accurate evaluation of the expected or normal variation in machine performance and the expected or normal variation in operator performance a basis is established for detecting the source of the difficulty. These evaluations of variability of performance may also offer clues to appropriate solution.

Description of the Relationship between Performances. As already indicated, human factors in industry are not to be studied in isolation but as they interrelate and interact among themselves and with the physical factors operating in the working environment. It is then obvious that the relationships between various factors must be described and evaluated quantitatively.

The relationship between two given variables will vary in degree or extent. For example, the correlation between variations in humidity and in workers' performance in a textile mill is high due to the extremely humid conditions that must be maintained in the weaving rooms. The relationship between these two variables in an installation where there is a relatively dry climate is low. Again, the relationship between finger dexterity and job performance in repairing jewelry is high, whereas this ability has a low relationship with the performance of power-shovel operators. A measure is needed, therefore, which will reflect the extent of relationship between different variables.

One statistical constant frequently used to describe a relationship is called the coefficient of correlation. The size of the coefficient varies with the extent of the relation. When the two variables under consideration are highly related the coefficient will have a high value. When the degree of relationship is low the size of the coefficient will be low. The exact meanings to be assigned the coefficient of correlation when it is used to describe variables in the industrial situation are discussed in later chapters.

Evaluation of Measuring Instruments. Every science is called upon to maintain a close check upon the instruments that are used in measuring its variables. Accurate description is achieved only if the measuring instruments are consistent in their performance and are measuring the particular factors which are under study. Evaluation of its measuring devices is particularly important for the science of psychology because of the multiplicity of determinants which underlie human behavior. So many factors operate to influence job performance and these factors vary so

widely on different occasions and with different individuals that a continuous check of measuring instruments is mandatory.

The valuation of measuring instruments reduces to two general forms. The first is concerned with the extent to which the instrument accurately repeats itself. This is called the reliability of the instrument. The second concerns the extent to which the instrument achieves the objectives set down for it, that is, the extent to which it does what it is designed to do. This is referred to as the validity of the instrument. A brief description of these two concepts will be given here inasmuch as they will recur quite frequently in the chapters to follow.

Reliability of Measurement. Reliability refers to the extent to which the measurement is repeatable. If the results are duplicated on successive occasions the instrument is considered reliable. Suppose that a test of hand-finger dexterity is given to a group of applicants for a job requiring a high level of skill in the use of the fingers and hands. Now suppose one month later the test is again given to the same group and that in the meantime the individuals have had no particular practice in the muscular coordinations being tested. Under these conditions, if the test were perfectly reliable, the rankings of the several applicants would be exactly the same on the two testings. That is, the individual getting the best score on the first testing would get the best score on the second, the individual getting next to the best score on the first testing would get next to the best score on the second, and so on down to the individual giving the poorest performance. As can be seen, this is a relationship between two sets of performance scores and therefore can be measured and represented by the coefficient of correlation. Such a coefficient is given the special name of the reliability coefficient.

It is not to be expected that the instruments measuring human factors will have perfect reliability. Human behavior is too complexly determined to realize such an ideal. It is expected, however, that the reliability will be high. Past experience dictates that in the measurement of simple functions such as reaction time, steadiness, simple arithmetic computations, etc., very high reliability coefficients can be obtained. In more complex functions such as reasoning, memorizing, complex judgment, etc., fairly high reliability coefficients can be obtained.

Validity of Measurement. Validity is basic to prediction. If an instrument accurately predicts the particular variable for which it was originally designed, then it is a valid instrument.

Like reliability, validity is a relationship between two sets of values. One set is the values obtained on the measuring instrument itself; the other set comprises the values on the variable to be predicted. In the example of the finger-dexterity test it was suggested that the test be given to applicants for a job which required a high level of hand-finger

dexterity. The problem is to select from the applicants those workers with the highest skill, and the argument is advanced that the dexterity test will differentiate the applicants in terms of this skill. The validity is expressed as the relationship between scores on the test and degrees of success on the job. It can be measured by the coefficient of correlation. This special form of the correlation coefficient is called the validity coefficient.

The validity of a measuring instrument is a function of the use that is made of it. If an instrument is used for a purpose for which it is unsuited, the validity will be low. If the functions measured by the instrument are the same as those to be predicted, the validity will be high. Validity then will vary from zero to very high depending on the closeness of similarity between the determinants in the two situations.

Instruments measuring human factors vary widely in their validity. When predicting proficiency for a job in terms of performance on an analogous test—an analogous test utilizes the job responses with little modification—the validity can be expected to be high. When predicting job success from aptitude tests—aptitude tests measure potential ability rather than present ability—the validity can be expected to be lower in value. Regardless of the nature of the problem, evaluations of the validity of the instruments of measurement are mandatory. Not to know the validity of the measures is to be completely ignorant of what is being measured.

ORGANIZATION OF THE REMAINING CHAPTERS

A brief description is here given of the organization of the subject matter of the remaining chapters. In the next three chapters the reader is made acquainted with the problems that must be solved in gaining an understanding of the two major components of work, namely, the job and the worker. Chapter 2 describes the procedures and techniques for learning about the demands of jobs and the qualifications that workers must possess in order to meet these demands. Chapter 3 deals with problems connected with measuring job proficiency. Inasmuch as the physical and human factors of work contribute to the effectiveness of the job performance, it is important at the outset that the reader be acquainted with the measurement of job proficiency. One of the standard procedures for measuring human performance involves the use of rating methods. These are discussed in Chap. 4.

The next group of problems concerns the selection and classification of workers. Chapter 5 outlines the important issues and points out the interrelations existing among the major aspects of placing workers on jobs, viz., selection, classification, transfer, and promotion. In these areas

the psychologist is again faced with the task of measurement. Standard procedures used in interviewing and personal-data analysis are presented in Chap. 6. The development of psychological tests and their function in selection and placement are discussed in Chap. 7. The widespread utilization of psychological tests in industry has provided a great deal of information on the reliability and validity of various kinds of tests as they have been used in the measurement of aptitude for a variety of jobs. Chapter 8 presents a review of these findings.

The next four chapters deal with the nature of work and some of the factors which condition the effectiveness of the worker on the job. Chapter 9 describes the major characteristics of work, including the concepts of fatigue, efficiency, mental work, and boredom. Conditions of work and productivity are treated in Chap. 10. Here are discussed such topics as measures of productivity, diurnal changes in output, rest pauses, illumination, and noise. Work methods and equipment design as they influence the worker's productivity are treated in Chap. 11. Problems connected with time-and-motion analysis, analysis and evaluation of movements, and arrangement of the workplace are considered. Chapter 12 describes accidents and the safety problem. The tremendous loss in manpower resulting from accidents raises a primary problem that continues to plague the best minds in industry.

The last group of chapters concentrates more specifically on the human factors in industrial production. Chapter 13 presents many problems incident to training and adjusting the worker to his work and shows how the worker's job adjustment encompasses many important factors which will be omitted if industrial training is limited to training only the specific responses involved on the job. Conditions influencing worker motivation are discussed in Chap. 14. Consideration is given to methods for studying motivation and to the dynamics of motivation and frustration. In the final chapter, Chap. 15, the worker is considered as a member of a group. The nature and characteristics of groups are described together with problems arising in the areas of supervision and leadership.

CHAPTER 2

Analysis of Jobs

A knowledge of the nature and requirements of jobs is a fundamental prerequisite for an intelligent attack upon all personnel problems in any organization, whether large or small, public or private. Unless a personnel officer has adequate knowledge of the nature and requirements of the job for which he is interviewing an applicant, he is not in a position to obtain the needed personal data from him. Without information concerning the interrelationships and interchangeability of jobs, it is impossible to make satisfactory transfers and promotions of workers. The most profitable beginning in the development of more effective methods of work is a study of the methods currently used on the job. A decision concerning the optimal shape for the handle of a screw driver can only be made when there is accurate knowledge of the specific kinds of tasks in which the screw driver is to be used. Labor legislation and labor agreements can be meaningful only if the jobs involved have been clearly defined. One of the first steps in attacking any employment problem, then, is to make an adequate job and worker analysis.

THE MEANING OF JOB ANALYSES

Definition of Terms. The terms *job* and *occupation* have been given a variety of different meanings. Sometimes they are used as synonyms; sometimes they are used to refer to different levels of grouping of industrial tasks. The most nearly standard terminology is that employed by Shartle.[15] This writer uses the following definitions of terms:

Position: A position is a group of industrial tasks that are performed by one person. In any organization there will be as many positions as workers when all positions are filled. In an office, for example, there might be four positions of clerk-typist, each of which would be filled by a different individual.

Job: A job is a group of similar positions. Thus in the above case the four positions of clerk-typist would constitute the job of clerk-typist in the organization. In any establishment there may be only one person or there may be many persons engaged in the same job.

Occupation: An occupation is a number of allied jobs. Thus one speaks of the jobs of clerk-typist, comptometer operator, and bookkeeper falling within the clerical occupation.

The terms *position, job,* and *occupation* shall here be employed in the foregoing meanings. The term *job analysis,* consequently, will refer to the procedures by which information is obtained concerning a specific job. The term *occupational analysis* will refer to the study of the relationships among the requirements of related jobs, and to the characteristics which the jobs within a given occupation have in common.

Job Analysis and Job Specification. The term *job analysis* refers to the procedures employed in collecting information concerning the nature and conditions of work involved in a given job. Job analysis is a descriptive process. It is the process by means of which a description is developed of the manner in which workers on the job actually perform the work, and under what actual conditions the work is done. The results of a job analysis are set down in the *job description.*

The job description must be distinguished from the *job specifications* which are formal statements about the nature and conditions of the work which is supposed to be performed. The job specifications are given by fiat or fixed by custom and tradition. They are the characteristics of the job that have been set down by law or prepared according to a plan before the actual work began, or the characteristics that have resulted from such factors as technological improvements, changes in the goods or services produced, the development of the organization, and collective bargaining conferences.

Both job description and job specification, then, are concerned with the methods and procedures of work, and the administrative and physical conditions of work. The job description tells how the job is really performed, and the job specifications indicate how management expects the job to be performed.

The Scope of Job Analysis. A job analysis will yield information such as the following:

A. Job name or title
B. Present methods and procedures of work
   1. Duties and tasks the worker performs
   2. Materials, supplies, etc., that the worker uses
   3. Tools, machines, and equipment the worker uses; description of the workplace
   4. Methods and procedures employed in carrying out the duties
   5. Responsibilities
   6. Supervision received and given
   7. Standards for output
C. Physical conditions of the work environment
   1. Place of work; inside or outside
   2. Conditions of illumination

3. Conditions of ventilation
4. Conditions of noise
5. Dangerous and unhealthful aspects

D. Relation of the job to other jobs
1. Helpers and assistants
2. Coworkers
3. Coordination of the work of the job with that of other jobs

E. Conditions of employment
1. Methods of selection of employees
   a. Techniques used, such as tests, interviews, etc.
   b. Promotion or transfer from other specified jobs
2. Length and hours of work periods
3. Amounts and methods of pay
4. Permanency or seasonality of the job
5. Opportunities for promotion and advancement

WORKER ANALYSIS

For many purposes it is desirable to list the qualifications that an individual must possess if he is to perform a given job in an adequate fashion. Although certain qualifications of the worker may be inferred from the job description or specifications, they do not actually define the worker requirements in any real sense. It is the function of *worker analysis* to describe the personal requirements of the worker. The results of such an analysis will be embodied in a *worker description.*

Worker specifications are the requirements that management sets for the job, and which the individual must meet if he is to be hired. The worker specifications for a job may or may not agree with the worker description. In many cases the specifications are too rigid or too loose, and frequently they are absurd.[20]

The Scope of Worker Analysis. A worker analysis will yield information such as the following:

I. Physical characteristics of the worker
1. Health
2. Strength and endurance
3. Body size
4. Handedness
5. Permissible physical limitations

II. Psychological characteristics of the worker
1. Sensory acuity
2. Skills and abilities, such as manual dexterity, ability to do arithmetic problems, etc.
3. Traits of temperament and character, such as honesty, emotional stability, etc.

III. Background of the worker
1. General education
2. Previous job experience
3. In-service training

THE NEED FOR SYSTEMATIC JOB AND WORKER ANALYSIS

Lack of Knowledge Concerning Jobs. It is the rare personnel officer, executive, or foreman who openly professes ignorance of the nature of the work carried on by those working under his direction. Nevertheless, in practice, superior officers do not always have adequate knowledge of the jobs under their direction, even though they may come in daily contact with the men working at them.

There are many factors at work, besides the lack of training in job-analysis techniques or even pure indifference, that may bring about this situation. A given job does not remain constant from year to year. Over a period of time the product or service may be changed, thereby necessitating changes in the original procedures and conditions of work. Interested and ambitious employees may enlarge upon their assigned duties. Inferior ones may lose portions of their work to others, or simply let them go. In other cases, although definite duties may be assigned, the employees may unofficially interchange some of them. In instances of undermanning or overmanning, the actual work may of necessity differ from some planned arrangement. It is important, then, for supervisors to be aware of all these factors if they are to maintain adequate knowledge of the requirements of jobs under their direction.

A study by Charters and Whitley well illustrates how incomplete may be the knowledge of an apparently clear-cut job.[1] These investigators set out to determine the different tasks performed by secretaries to businessmen and executives. The members of their research staff together with six secretaries recorded all the secretarial duties they could recall. They succeeded in compiling a list of 166 duties. In order to ascertain the completeness of this list, a chart was prepared covering a working week, each day being divided into time periods. Copies of this chart were distributed to 125 secretaries, who recorded their actual activities as they performed them in each time period during the day. By this means, errors due to forgetting were reduced to a minimum, and an almost complete coverage of the actual duties of secretaries was assured. A tabulation of the duties from these charts revealed that 871 different duties were mentioned. This was over five times the number in the original list.

It is to be noted that these experienced investigators, working with six persons actually engaged on a given job, were able to compile only a partial picture of that job. It seems highly improbable, then, that the executive who, as a small part of his regular work, hires and fires his secretary could have an adequate knowledge of the tasks his secretary is called upon to perform.

The Need for Uniformity of Job and Worker Descriptions. Systematic job and worker analyses are necessary in order to provide a common language in the employment field. In different organizations or firms, jobs involving exactly the same tasks may have different titles or names. On the contrary, jobs involving quite different tasks may have the same title. For example, a worker who operates the machine that starches and presses cloth in preparation for dyeing or printing is known variously as a mangle tender, cloth-finishing-machine operator, mangle ranger, or trojan ironer. On the other hand, the job title *mounter* is used in the clothing industry to refer to a laborer who collects bits of waste material, in the furniture manufacturing industry to a semiskilled worker who attaches carvings and moldings to various parts of furniture, and in the optical industry to a skilled craftsman who assembles and repairs glasses.

In an effort to standardize terminology, the U.S. Employment Service analyzed a large number of different jobs and has published a "Dictionary of Occupational Titles" which gives definitions and summary descriptions of some forty thousand jobs.[19] This dictionary has had a profound effect on personnel practices by providing a basis for a more thorough understanding of the nature of jobs and their interrelationships.

Importance of Systematic Job and Worker Analysis. Systematic procedures in job and worker analysis are a necessity if reliable and valid information is to be obtained. The hit-or-miss methods which commonly are employed in obtaining information about jobs and their demands upon workers do not give satisfactory results. The information obtained by such casual observation will be far from complete and almost certainly will contain serious inaccuracies.

For a surprisingly large number of jobs no descriptions or specifications of any kind are available. In others, specifications consist of only a sentence or two outlining major areas of duties and responsibilities. In many instances a strange faith is put in the job title as conveying all the information needed concerning the job. Thus "everyone knows" that the major duty of a college professor is teaching. Yet he may spend more time dealing with administrative matters and budgets than he spends in teaching.

A good example of the confusion that results from a lack of systematic job studies is well illustrated by the experience of a milk distributing company.[3] The executives of this firm disagreed on the function of their deliverymen, some considering them as laborers and others as salesmen. An analysis of what the men actually were doing showed that this disagreement and disorganized thinking on the part of management were reflected in the performance of the employees. At the one extreme, about one-fourth of the deliverymen believed their work to be a simple rou-

tine job and occupied themselves solely with deliveries. At the other extreme, about one-third conceived of their jobs as primarily selling, and therefore in addition to their delivery work vigorously solicited further orders. Since the observations of actual solicitations revealed that some 80 per cent of them resulted in additional sales and that the calls were well received by the customers, the conclusion was drawn that the job of deliveryman was clearly a sales position.

PURPOSES FOR WHICH JOB AND WORKER ANALYSES ARE MADE

The Development of Measures of Job Proficiency. In order to develop adequate indices of success on a job it is necessary to take into account all the elements of the work involved. It would be inaccurate to describe a rapid typist as a good one if she makes many errors and turns out copy that is not neat. Rapidity of typing is a measure of job success, but it is only one among several possible measures and does not provide a complete description of success. Reference to the job specifications would indicate that both speed and accuracy are important.

It has already been pointed out that without a systematic job analysis a complete picture of a job cannot be obtained. Consequently, any measures used to evaluate the performance of employees should be worked out from a systematically determined set of job specifications.

The Functional Organization of Jobs. In practically every governmental or industrial organization several persons deal with a given item, each processing it in relation to a given objective. For effective operation, work must pass through definite channels, and levels or points of authority and responsibility must be definitely established. A simple case in point is a machine shop where the work flows from one employee to another. An order is received by a clerk, who turns it over to a shop foreman. The foreman checks the specifications and, in turn, passes the order to a machinist. The latter plans the work, obtains the necessary materials from a stock-room clerk, who previously has received them through the purchasing office, and then proceeds to carry out the job. The finished product is sent to a shipping clerk, and the order, plans, and specifications finally are turned over to a billing clerk. In many organizations well-defined routes have not been laid out. Confusion then results when the current haphazard organization is faced with an unusual situation, such as the absence of a key employee, or a somewhat different type of work.

Job Evaluation. The setting up of accurate wage scales for different jobs can be done only through job descriptions and specifications. In an office, for example, the employees will be engaged in a wide variety of activities. Workers will be indexing, filing, typing, handling corre-

spondence, distributing mail, keeping books, operating business machines, making up payrolls, supervising other workers, etc., and the question of what wage should be paid for each type of work must be answered. Each job must be evaluated in terms of its importance, its difficulty, the degree of supervision required, the amount of experience or training necessary, and similar factors. Equal pay is then assigned to equal work. Obviously, detailed specifications of the jobs must be available before they can be satisfactorily evaluated.

The Development of Techniques for Use in the Selection of New Employees. Job and worker analyses are the first step in the development of selective techniques. They are the best means of discovering the traits and personal characteristics leading to success or failure on the job. A good set of job and worker specifications are a valuable source of clues concerning the kinds of questions to be used in the employment interview and the application blank, and the types of tests that will probably be most effective in the selection of new employees.

The Development of More Effective Methods and Procedures of Work. To improve the effectiveness of methods and procedures of work one should begin by obtaining an adequate description of the current processes of work. From an account of the movements and operations involved in carrying out the work, points needing improvement may readily be detected. The degree of success to be expected from changing the procedures of a specified task will be a function of the adequacy with which the original procedures have been described.

The Development of More Effective Tools and Equipment. Job and worker analyses have been almost completely neglected in the field of the design of tools and equipment, where they should play an important role. At the present time, tools and equipment, by and large, are designed by rule-of-thumb methods or a priori analyses. The case of the common heavy screw driver will serve as a good example.

In the few studies made of this type of tool, effectiveness has been evaluated in terms of the amount of force that can be exerted. An adequate analysis of the actual jobs where heavy screw drivers are used probably would reveal that the occasions where screws and bolts are turned by means of a screw driver against a constant strong resistance are relatively few. In many cases a heavy screw driver is used simply because the screw or bolt is a large one and not because there is any great resistance to its turning. In others the initial resistance impeding the turning of the screw or bolt is slight and increases as it is driven home. Analysis of jobs where screw drivers are used would indicate the specific characteristics of the various kinds of situations in which screw drivers are needed, and thereby provide a much better basis for the design and improvement of this tool. It seems likely that tools evolved

by such a process would differ in many unexpected ways from tools found on workbenches today.

The Development of Training Programs. The two main problems in the setting up of training programs for employees are the determination of what to teach and how to teach. Through job and worker analyses, information on these points can be obtained. Such analyses will reveal the skills and knowledge it will be necessary to develop in the training program. With information at hand on the nature of the different aspects of the work, training for the job can be divided into units. These units can then be arranged in optimal instructional order according to their difficulty and the logical relationships existing between them. In addition, through such analyses the problems of transfer and promotion will be better understood, so that the training program can prepare the workers for a change of jobs on the same level or for a change to higher level jobs involving more responsibility.

SOURCES OF INFORMATION IN JOB AND WORKER ANALYSES

There are a number of different sources to which the job analyst can turn for information concerning the job. The more sources he draws upon the more valid and complete will be the information he obtains. Information obtained from one source can be checked for dependability and completeness against that obtained from other sources.

Observation of Workers. The source of information probably most widely used is simple observation of the worker as he is engaged in carrying out the normal activities involved in his job. When these activities are slow enough the analyst can make his observations directly. When they are too rapid to permit accurate personal observation and recording, photographic methods may be employed.

For some purposes and for some types of jobs, simple observation of the worker does not yield adequate information. With administrative positions, for example, it is unlikely that from observation alone the analyst can get accurate knowledge of the activities of the administrator, since so much of his activity is "headwork," consisting of making plans, setting policies, etc.

Interviews. Interview procedures may be used to collect information about jobs, but they are subject to misuse. Misinformation may result if the interviews are conducted uncritically. It would seem obvious that the person who should know most about the job is the person doing it. However, there are several important difficulties encountered in using an interview with the worker as a source of information. In the first place, the success of the interview will be dependent upon the degree of cooperation the analyst is able to obtain from the worker. If the

worker feels that the study is being made for some purpose that ultimately will be to his disadvantage, he will not be willing to give a complete or accurate picture. For instance, he will not cooperate if he believes the interview may eventually result in the instituting of speed-up methods. Management's use of interviews with employees, then, is intimately connected with the problem of employee motivation and morale. Once rapport is established, however, valuable information can be obtained since most workers are fairly well identified with their jobs and are interested in discussing problems related to them.

Even if rapport is established, there are still other difficulties with this source of information. The employee might forget to mention certain details of the work. Duties and responsibilities seldom performed, even though important, may be overlooked. He may be willing to talk but might be unable to put into words a description of certain of the work processes. Many manual workers and frequently many office workers have not developed or do not have opportunity to develop any great facility in the use of words for purposes of description. In the interview they cannot always be expected to give accurate word pictures of what they do on the job.

Interviews of associates and superiors frequently are not particularly valuable as a primary source of information. Only occasionally will such associates and supervisors have intimate knowledge of the job of another, even though that person may be working under their direction. Consequently interviews with them should be used chiefly for checking on information derived from other sources.

Questionnaire Responses. The use of questionnaires with workers or with their associates or superiors involves the same difficulties as are found in the interview. The adequacy of this method, furthermore, depends upon the proper wording of the questions. In an interview, if a question is misinterpreted or not understood, the job analyst can repeat it in other words. With a printed questionnaire, filled out when the analyst is not present, such interpretation or amplification cannot be made, and errors are likely to go undetected.

Materials of Work. Knowledge of the tools, equipment, machines, printed forms, etc., used on a job, is a useful secondary source of information in job and worker analyses, serving to check the information obtained by other means. Such knowledge also may aid the analyst in formulating questions when the interview or questionnaire method is used. For some purposes knowledge of the materials of work might furnish all the information required.

Published Manuals, Bulletins, etc. Frequently overlooked sources of information concerning a job are the previous studies that have been made of it. The amount of occupational information available is far

greater than is generally envisaged.[15] From available published information the job analyst often is able to obtain data that may be useful as tentative or partial job descriptions, and certainly will be helpful in planning his study. In most instances, information obtained by this means will again be used chiefly to supplement that obtained from other sources.

Performance of the Job. For many jobs the analyst gains insight by actually performing the tasks required by the job. Neither observation of the worker performing a given activity nor a description of that activity by a worker always gives a complete and accurate conception of the work involved. From the "looks" of an activity one can never get exactly the "feel" of it. The surprise of a nongolfer who finally is induced to "take a swing at the ball" is a common tale. A job or worker analysis will be much more complete and accurate if the analyst has an opportunity actually to try his hand at the job. Skills and abilities not revealed through other sources will become apparent.

TECHNIQUES IN JOB ANALYSIS

Some of the information sought in job analysis can be obtained without special techniques. Simple observation will reveal whether the work is done inside or outside. Reference to the personnel records will reveal the hours of work, amounts of pay, etc. The chief problem is to obtain a valid and complete account of the duties the employees are expected to perform and the methods used in carrying them out. The following discussion of job analysis will be confined to techniques for obtaining such descriptions.

The techniques used in analyzing a job and the manner in which the activities of the workers are described necessarily will be a function of the type of job being studied. The activities of a factory worker, for example, are almost wholly manual, however complex the job may be. On the other hand, the manual activity of a clerical worker or salesperson is less important. The interest in the latter case will be centered primarily on the duties performed. More specifically, information will be sought concerning the types of situations with which the employee is confronted, and the procedures he uses in adapting to them. In the case of a factory worker the interest will be focused primarily on the movements the employee executes in carrying on his work.

The Analysis of Movements. Several techniques are available to aid the analyst in getting a description of the worker's movements when these involve responses of very short duration. The movements made by a mail clerk in the simple job of folding bills and inserting them in envelopes will be used to illustrate the problems arising in connection with the analysis of movements, and the techniques developed to study them.

The total time required to pick up a bill, fold it, and insert it into an envelope is about 14.25 sec. A job analyst observing a clerk carrying out this job would find it difficult to write down a full description of each of the movements. They are so rapid that by the time a description of the first movement was written down the worker would have completed that unit of work and be well along on another.

One solution to this problem is to have a system of shorthand symbols that the analyst can jot down as he observes the worker in action. Some years ago Gilbreth realized this need and attempted to classify all movements employed in industrial work into 17 basic types, providing a symbol for each.[8] These elements of movements, called *therbligs*, are shown with their notations in Table 2-1. Other symbols or abbreviations are

Table 2-1. Gilbreth's 17 Elements of Movement or Therbligs and Their Symbols

Search	Inspect
Find	Pre-position, prepare for next operation
Select	Release load
Grasp	Transport empty
Position	Wait, unavoidable delay
Transport loaded	Wait, avoidable delay
Assemble	Rest for overcoming fatigue
Use	Plan
Disassemble, take apart	

often used, the number of symbols and the particular notations varying greatly with different analysts. When a very exact description of the movements is desired, a rather detailed system of basic movements such as Gilbreth's is used. For a more generalized description of a worker's movements, four or five basic symbols may suffice.

The analysis of a total sequence of acts into its constituent movements has been termed *motion study*. Some time before Gilbreth's work, Taylor had made studies of the time taken to perform each particular operation in an industrial task. These were called *time studies*.[17] Today the two types of study usually are combined. The resulting *time-and-motion* analysis shows not only which movements are being employed but also the time required to perform each of them. In making a *time-and-motion*

study the job analyst simply notes a therblig on the record sheet every time the worker makes a new movement, and times each of these movements by means of a stop watch.

The results of a time-and-motion study are ordinarily presented in a *motion-cycle* chart, such as that shown in Table 2-2, for the movements

Table 2-2. Analysis of the Movements Made by a Mail Clerk in Folding a Bill and Inserting It into an Envelope

Left Hand

Hand from pile of folded envelopes to resting position. 2.00 sec.

Wait. 2.00 sec.

Hand to top of folded bill. 0.75 sec.

Hold bill folded. 2.00 sec.

Reach for envelope. 1.75 sec.

Select, grasp and carry envelope to working position. 2.00 sec.

Hold envelope for inserting. 2.00 sec.

Carry filled envelope to dump pile and release. 1.75 sec.

Time in ¼ sec.

Right Hand

Select and grasp bill and bring to working position. 2.25 sec.

Near side of filler folded over and edge matched with edge of far side. 2.00 sec.

Hold bill folded. 0.75 sec.

Hand to edge of fold. 0.75 sec.

Smooth fold. 1.00 sec.

Grasp bill and wait. 2.25 sec.

Carry bill to working position. 1.50 sec.

Insert bill in envelope. 2.00 sec.

Reach for another bill. 1.75 sec.

Repeat the above cycle of movements

involved in the bill-inserting task. The advantages of such a chart over a running verbal description of the movements are obvious. Not only are the individual movements themselves easier to detect, but also their pattern or sequence and their relative durations can be appreciated more easily.

Sometimes the analyst utilizes motion pictures for recording fast movements. The speed and variety of movements involved in some jobs are so great that an observer cannot possibly record all the worker's movements, much less time them with a stop watch. To overcome this diffi-

culty motion pictures may be taken of the worker as he carries on his activities. When the film is projected at slow motion it is a simple task to note the different movements and to jot down the corresponding therbligs. By placing a clock near the worker being photographed the exact time required for each movement can also be determined from the film.

For some manual jobs the use of motion pictures is especially valuable, since the paths in space through which the worker's hands move can be

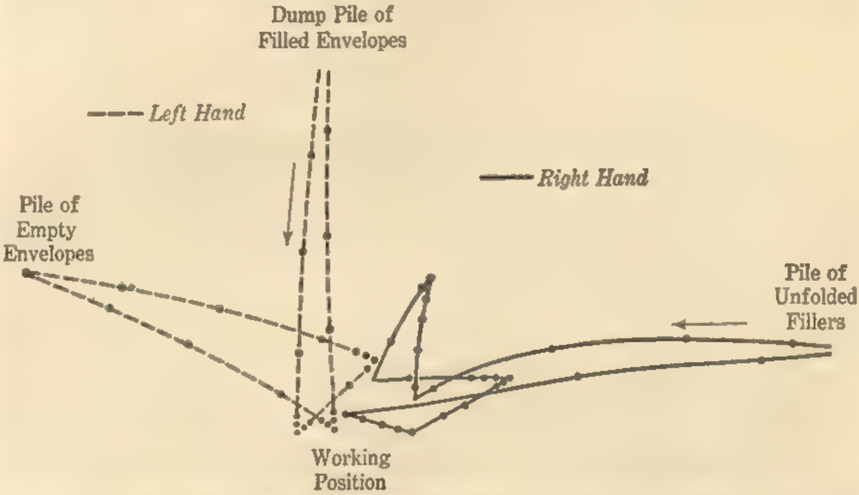


FIG. 2-1. Micromotion chart of movements made by a clerk in inserting bills into envelopes.

reconstructed. Depending upon the rapidity of the movements, each second, fourth, etc., frame of the film is projected upon a large piece of paper, and the position of each hand is marked on it by means of a dot. When the position of the hands has thus been indicated throughout a given cycle of work, the dots are connected. Figure 2-1 is a chart constructed by this method, showing the movements made by a mail clerk in folding and inserting bills.

Another way of obtaining a record of the paths of the hands is by means of the so-called *cyclegraph*. A small electric light is attached to the hands or other moving parts of the worker, and a photographic plate or film is exposed while he works. As the worker moves his hands a picture of the path of the light is made on the plate. The picture is similar in nature to the one shown in Fig. 2-1. In order to time the movements an interrupter is introduced into the light circuit. By having the light flash on more suddenly than it flashes off, the photographic plate will show a series of pear-shaped dots from which the direction of the

movement can easily be ascertained. The use of either of the photographic methods discussed above has been termed *micromotion study*, and the resulting chart, a *micromotion chart*.

Photographs of a worker can be taken by two cameras placed at different angles, thus giving a three-dimensional record of the movements. By using stiff wires a model of the movements can be made. These models showing the total act have been found especially useful in teaching new operators. Movements represented in three dimensions are much easier to grasp than those depicted on a two-dimensional chart, demonstrated by another worker, or simply described verbally. The novice has an opportunity to perceive the total sequence of movements at one time and thus can better understand the interrelationships of the different movements.

The Analysis of Duties, Situations, and Procedures. The varied nature of the work involved in nonmanual jobs makes it impossible to set down specific techniques that can be applied to all of them. With such jobs the analyst is interested in knowing the types of situations that confront the worker, and the procedures he utilizes in dealing with these situations. The problem is to obtain complete and valid information. When dealing with positions in a highly organized occupation it is a relatively easy task to obtain satisfactory information and to make sure that it is complete, since the lines of work are already laid down in definite channels, sequences, and groupings. On the other hand, when the analyst is confronted with a job the duties of which are not clearly prescribed, the task is extremely difficult, since there are no obvious points of departure for making the analysis or any organization or sequences to follow in the analysis.

In the case of a job like that of police officer the analysis is greatly facilitated by the highly structured character of the police service and the systematization of its activities. This organization of its activities has evolved from a long history of having to cope with the rigorous demands of legal problems. The analyst will find jobs classified into divisions, such as the patrol division, traffic division, and crime-prevention division, each with clearly defined objectives. Within any division jobs are readily classified into major assignments. In the traffic division, for example, men are assigned to traffic control, parking control, motorcycle control, etc., each assignment having prescribed responsibilities. All that remains for the analyst to do is to discover the specific situations encountered in each assignment and the procedures followed. This can readily be done because in most instances the individual officer records or reports to his superior the nature of his activities during every tour of duty.

Compared with this highly structured situation is the job of sanitary inspector. While the general objectives of this job are prescribed, the

manner in which these objectives are achieved varies markedly from one health department to another and from one inspector to another. This arises in part because of the lack of standardized training requirements for inspectors, and in part because the job is a complex one with some technical phases and some phases bordering on law enforcement. The task for the analyst with a job such as this would be to discover all the situations in which any inspector is involved. The activities then would be classified as to general and specific areas. The final description of the job probably would not represent the work carried on by any particular inspector, but rather would give a picture of the total activities of the job.

Charters and Whitley's study of secretaries is especially interesting from the point of view of the methodology.[1] As was previously pointed out, these investigators were not satisfied with their own efforts to list all the different secretarial duties. Consequently, in seeking a complete and valid list of duties they prepared a daily schedule chart, which was filled out by a group of cooperating secretaries as they actually carried out their work. This procedure assured the completeness of the list inasmuch as the opportunity to forget details was virtually eliminated, and it also enhanced its validity since the secretaries wrote down the duties they *actually* performed and not those they *thought* they did or should do.

The procedure of noting down functions as they are actually performed may be very revealing. Many unsuspected aspects of a job may be uncovered. For example, in a study of salesgirls in a notions store it was found that one-third of their time was spent in activities not ordinarily connected with the job of salesperson. These were acting as messenger, keeping personnel and stock records, and performing general service functions. In a study of persons engaged in the construction of tests for civil service examinations it was found that they were able to pursue their primary duty of making up tests for less than half of their working hours. The remainder of their time was spent in administrative activities, making reports, and attending meetings and conferences.

Under certain conditions it is not desirable for the workers themselves to report on their own activities. They may become so engrossed in their activities that they forget to note them down. In addition, taking time out to record activities might seriously interfere with them. It would be absurd to expect an airplane pilot to list his activities at the time he is landing his craft. Thus the recording of activities by an observer may be necessary.

In some jobs, such as that of foreman or salesman, the analyst need not observe the worker's activity continuously. Instead, he makes his observations at regularly spaced and predetermined intervals of time.[2] Depending upon the rate of activity, the analyst might make an observa-

tion every 10 sec., 1 min., or 5 min., and record the activity at the moment of observation. This technique, which has been termed *time sampling of activity,* has the advantage of permitting the analyst to collect information simultaneously from a number of different workers. For example, individuals engaged in some group activity, such as riveter and bucker-upper, or airplane pilot, copilot, and navigator, can be observed in rotation. The major disadvantage of this technique lies in the fact that activities occurring infrequently or reactions to emergency situations may not occur during a period of observation and therefore may not be noted. However, in the few investigations where the time sampling technique has been reported, highly consistent results have been obtained.

JOB SPECIFICATION

One of the major personnel problems that has received little attention is that of setting job specifications. While, to be sure, duties and responsibilities are assigned to jobs, there has been relatively little thought given to the development of systematic approaches to the allocation of duties and responsibilities. The usual approach seems to be quite intuitive, those confronted with the problem "playing by ear" rather than following prescribed procedures or even general principles. In an office, for example, filing duties might be assigned to an accounts clerk and the placing and receiving of telephone calls to a stenographer because the former is "more familiar with details" and the latter is "more accustomed to dealing with people."

Usual Objectives in Setting Job Specifications. In the setting of job specifications there appear to be two general objectives and a number of specific ones. The first general objective is to ensure that all required functions in an organization are assigned to one job or another. Thus, on an assembly line, the worker who adds the last part might also be charged with the inspection of the total assembly. The second objective is to maximize the flow of work and thereby minimize stoppages and delays. In order to achieve this objective a number of jobs may have to be considered simultaneously. If too many activities are assigned to one job, the person filling it may have too great a burden and hence slow down the flow of work in the organization. In addition, there are specific objectives to achieve such as allocating activities in order to minimize hazards, provide independent checks on operations, and reduce wastage.

All the foregoing objectives have to do with the operations with which the organization is concerned. But there are other types of considerations which enter into the allocation of duties and responsibilities to jobs. Duties and responsibilities will be assigned to certain jobs rather than

to others in order to satisfy certain laws. For example, the lifting of heavy objects is assigned only to jobs filled by men, and the treatment of injuries only to physicians and nurses. If the available market contains mainly persons of little skill, complex tasks may be fractionated into parts in order to eliminate long training periods. In labor-management agreements allocation of duties and responsibilities enters as a factor. They may be involved in the so-called fringe benefits, or in jurisdictional disputes.

From the foregoing it is clear that the considerations that enter into the assigning of duties and responsibilities to jobs are many and varied. Furthermore, the objectives to be achieved by such allocation are equally varied. Until the important considerations are outlined and the objectives codified, little more than partial solutions can be expected.

Perception of the Job and Job Specification. Characteristics of a job are perceived as varying in importance. Since a job is an assemblage of tasks that are to be performed under certain given conditions, it is an intangible. To be sure, it acquires a sort of substance in that equipment, materials, desks, etc., are utilized in its performance. Furthermore, it may become associated with a spatial location, as a particular place in a shop or an office. But these are only partially defining characteristics. The duties, responsibilities, and functions of the job are likely to assume far greater importance in any given individual's perception of what the job is and what it involves.

Those who are concerned with a job differ in their ideas about its nature and characteristics. In many areas painters consider that for their job it is management's responsibility to supply necessary tools and equipment. Carpenters, however, have the opposite view, holding that for their job it is the responsibility of the employee. The head of a department in a retail store might perceive the job of salesperson as involving responsibility for arranging merchandise but not for cleaning behind the counters, whereas service personnel might perceive the job as involving both duties.

In many instances the origins of ideas concerning a job will be obscure. The perception of the job may be quite unclear, as in the case of the milk deliverymen cited earlier. Here it appears that the different ways in which the deliverymen viewed their job were largely the result of a lack of structuring of the situation on the part of management. In their study of the duties of secretaries, Charters and Whitley found wide discrepancies in the duties that secretaries performed. Analysis of the responses of 715 secretaries to a check-list questionnaire of the 871 duties indicated that only 3 per cent of the duties were performed by three-quarters or more of the secretaries. From this study it is apparent that

the job of secretary is not one with clearly or universally prescribed or perceived duties.

There is a dire need for systematic development of job specifications. Certainly the problems involved have not been systematically analyzed or precisely stated, and adequate techniques have not been developed. Few direct comparative studies have been made of the manner in which jobs are perceived by workers and management. Nevertheless, the ways in which jobs are viewed by the interested parties undoubtedly is of considerable importance. Divergent perceptions by supervisors and workers are a factor in labor-management disputes. Certain duties and responsibilities might be seen to be part of a job by management but not by labor, while others might be seen to be part of the job by labor but not by management. It would be expected that the chances of labor peace would increase as the perceptions of the jobs by both parties become more congruent. Similarly, jurisdictional disputes between unions might arise when there is a lack of congruence in perceptions. Each of two unions representing workers on nominally different jobs might perceive the same set of duties and responsibilities as part of the job held by workers it represents.

THE PROBLEM OF WORKER ANALYSIS

Aims of Worker Analysis. Among the many specific statements about the purposes of worker analysis it is possible to differentiate two general aims: the determination of minimum requirements of jobs, and the determination of the relative importance of various abilities, traits, and other worker characteristics for the performance of jobs. Implicit in the first aim are two presumptions. First, it is presumed that there is a well-established point on the continuum of job proficiency below which individuals are clearly inadequate and above which they are clearly adequate. Second, it is presumed that, for a worker characteristic with a minimum requirement greater than zero, all individuals, or at least most, who fall below this minimum also will fall below the critical point in job proficiency, while all, or at least most, who possess more than the minimum amount of the characteristic will perform on the job at a level above this critical point.

Consider an example from the transit industry. According to the first aim it must be possible to make a statement like the following: Bus drivers who have three or more accidents per month lose more money for the company than they earn. Then, if the minimum educational requirement for bus drivers is 12 years, it must be possible to show that most bus drivers who are high-school graduates incur fewer than three

accidents per month while most who have less than 12 years' education incur three or more accidents per month.

On at least two counts this first aim of worker analysis is not wholly satisfactory. To begin with, in most instances it will be difficult or impossible to set rigorous and meaningful critical points on the continuum of job proficiency. Any critical points will vary with such conditions, for example, as the nature of the labor market. If good workers are scarce a lower quality or amount of work might have to be considered acceptable. Few organizations have such an accurate cost-accounting procedure that they can determine that a worker who produces less than a given amount is a financial loss to the organization. Furthermore, as will be seen in the next chapter, proficiency in many jobs cannot be gauged in terms of production. A second difficulty is that it is unlikely that a point can be found on any worker requirement—such as education, capacity to accept responsibility, or finger dexterity—below which all or almost all inferior workers fall and above which all or almost all successful workers fall. Such a situation would require that variations in the worker requirement be perfectly correlated with success in performance of the job. Such a correlation has never been empirically demonstrated.

The second aim, that worker analysis is concerned with the discovery of the relative importance of various worker characteristics in job performance, is a more reasonable point of view. Here importance of a given characteristic would be expressed in terms of the extent to which individual differences in it are correlated with individual differences in job performance. Thus it might be said for a particular job that amount of education is less important than visual acuity because years of education are correlated with job proficiency only to the extent of .10, while measures of visual acuity are correlated to the extent of .30. With this approach there are no implications concerning any critical points to be found on the continuum of job proficiency, or the setting of minimum points on worker requirements for a job. The minimum job requirements are to be set on the basis of other considerations such as the level of job proficiency desired at the given time, the nature of the available labor market, or even sheer convenience.

Estimated versus Measured Worker Characteristics. There are two quite different ways of securing information concerning the importance of various worker requirements. A given job may be described in terms of the skills, abilities, and other characteristics it is believed the worker must have in order to perform the job successfully. This type of description may be said to be concerned with *estimated worker characteristics.* The personal characteristics deemed necessary in the worker are estimated from the analyst's knowledge of the nature of the work or from knowledge possessed by others. The job specifications or description or

similar information obtained from unsystematic job studies form the basis of estimated worker characteristics. The second type of description may be said to be concerned with *measured worker characteristics.* With this approach an attempt is made to obtain an objective description of the personal characteristics of individuals who are actually working on the job.

Both of these ways for describing worker requirements have certain advantages and disadvantages. The vagaries of personal judgment inherent in the estimation of requirements by analysts are avoided by the approach of measured characteristics. On the other hand, the approach of measured characteristics is costly both in time and money, and is subject to any limitations imposed by the particular sample of workers available for the study.

THE APPROACH OF ESTIMATED WORKER CHARACTERISTICS

General versus Specific Meanings in Worker Requirements. In the earliest worker analyses all that was done was to list those traits, abilities, skills, and other characteristics which, in the opinion of the analyst, were important in the performance of the job. Such lists usually suffered from the use of vague and abstract terms that rendered them useless for any precise evaluations. For instance, it might be said that the job of machinist required manual dexterity, judgment, concentration, and ability to plan. Impressive as this list of job requirements may seem, it is obvious that the same four traits also could be used to describe the requirements of the job of typist or of cook. This deficiency has been remedied to a considerable extent in more modern approaches by using specific rather than generalized traits, and by careful definition of terms. Specifying the traits with more exact definitive descriptions and exemplifying them with samples of actual behavior have improved the results.

Relating Job Requirements to Specific Job Behaviors. The important consideration for achieving specificity, and hence meaningfulness, of worker requirements is to relate the requirements to specific behaviors on the job. When Charters and Whitley sought to discover the traits important for the job of secretary they interviewed persons who supervised the work of secretaries.[1] They asked questions such as the following: "When hiring a secretary, what are the qualities in which you are especially interested?" "Who was the best secretary you ever had? In what respects did she differ from the others you had?" "Have you ever discharged a secretary? Why?" With these questions the actual performance on the job formed the basis for the respondents' answers. Vague, abstract, and assumed qualifications tended to be avoided.

The Critical Requirements Approach in Worker Analysis. Flanagan has pointed out that a listing of the duties and responsibilities connected with a job, even though the list seems complete, may leave out requirements which are of a critical nature.[7] For example, the primary duties of a bottle-capping-machine operator would be stated as "insert bottles one at a time into the capping machine, depress capping lever, remove bottle, and examine for security of the cap." A worker on this job would more likely be considered outstanding, however, if he seldom caused a work stoppage on the production line by permitting his machine to run out of caps than if he capped bottles at a very high rate. Similarly, a cashier whose duties are to "check sales slips and make appropriate change" would more likely be fired for stealing than for being slow in making change. It is clear that there are some kinds of behavior that are especially critical in the sense that they are much more likely to lead to administrative actions or judgments concerning the worker than is the goodness or poorness which characterizes his performance of the prescribed duties and responsibilities. In doing a worker analysis, then, special attention should be given to the discovery of critical job behaviors of workers.

Preparing a List of Critical Incidents. Critical incidents that occur in connection with an individual's job can be uncovered by interviewing the individual, his superiors, his subordinates, or his associates. Questions which might be asked of a superior are "Who was the last employee you had to have transferred from your department?" "What was the last straw that led you to take this action?" In the questioning, attention is focused upon actual behavior, and especially behavior that leads to important judgments about a real individual.

Questions are framed to reveal both particularly effective behavior and particularly ineffective behavior. An example of a critical incident revealing effective behavior on the part of a foreman as reported by his supervisor would be, "I was away when this rush order came through, so he couldn't check with me, but he saw right off how important the job was so he stuck his neck out and okayed the orders for labor and material." A critical incident exemplifying ineffective behavior of a machine operator would be, "He hardly ever knows when his machine is grinding too close and when he does he has to get somebody else to adjust it."

Incidents collected for a particular job can be grouped into major areas or types such as delegation of authority, planning of work schedules, initiation of new procedures, etc. The final result is a list of critical requirements for the job stated in terms of concrete behavior of workers rather than in terms of abstract traits. By focusing attention on the critical phases of the job, Flanagan's approach of critical requirements fur-

nishes a sound basis for the development of measures of job proficiency.

The Importance of Quantification of Worker Requirements. The procedure of simply listing essential traits is inadequate because it does not indicate the degree of importance of the traits. Two or more jobs might not differ in terms of the particular traits, abilities, and skills that are important, but they might differ radically in terms of the relative importance of these characteristics. For most purposes it is not sufficient to say that a given trait is or is not important. Rather it is necessary to know the degree of importance of the trait.

A watchmaker, an assembler of electrical equipment, and a cherry pitter all need to have more than an average amount of finger dexterity, but certainly not the same amount. It is obvious from the nature of the work that the watchmaker needs more finger dexterity than either of the other two workers, and that the assembler requires more than the cherry pitter. It would be useful to know the amount of the differences in finger dexterity required by these three jobs. It is necessary, then, that the worker analysis provide indications of the degree of importance of each of the traits deemed necessary for a job.

TECHNIQUES IN ESTIMATED WORKER ANALYSIS

A number of different techniques for worker analysis that yield quantitative estimations of the importance of various worker characteristics have been developed. The earliest systematic attack on estimated worker analysis was made by Viteles, and his method has largely set the pattern for all later developments.[21] As Viteles viewed the problem it appeared that what was needed was a list of all skills, abilities, and traits that might be important on any job, and a scale on which the analyst could rate the degree of importance of each of these characteristics for the job in question.

Viteles stressed the need for careful definition of terms. In his procedure and in all other good procedures each worker characteristic is carefully defined and specified through example.[5] Following is an illustrative definition of a worker characteristic:

Division of attention

Definition: Ability to perceive and respond to a quickly changing series of multiple and varied stimuli.

Illustration: Telephone-switchboard operator.

Summary of duties: plugs in, disconnects, and dials incoming calls, often taking down messages at the same time.

Job elements which justify a rating of high importance: A number of signals occur at the same time and calls come in continuously.

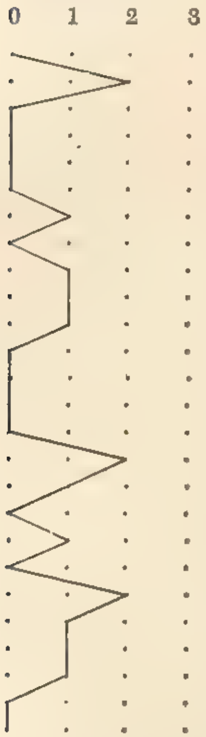
Other jobs for which division of attention is important: airport-control-tower operator, taxicab driver, traffic officer.

A typical worker-analysis schedule is shown in Fig. 2-2. Generally a *job profile* is made by drawing a line through the ratings, as indicated. The principal requirements of the job can then be readily noted.

WORKER ANALYSIS SCHEDULE

Job Title: Stenographer

Amount of Characteristic — *Characteristic*



1. Strength
2. Dexterity of fingers
3. Dexterity of hands and arms
4. Dexterity of foot and leg
5. Eye-hand coordination
6. Foot-eye-hand coordination
7. Perception of form, shape, size
8. Perception of speed of moving objects
9. Keenness of vision
10. Color discrimination
11. Keenness of hearing
12. Sense of smell
13. Sense of taste
14. Touch discrimination
15. Muscular discrimination
16. Memory
17. Numerical facility
18. Understanding of mechanical devices
19. Division of attention
20. Oral expression
21. Written expression
22. Ability to make decisions
23. Ability to plan
24. Initiative
25. Emotional stability
26. Ability to meet and deal with people

Description of Scale Values:

3. A very important characteristic, a capacity possessed only by about one fifth of the population, a high degree of ability

2. A moderately important characteristic required, a capacity possessed by more than half of the population, an average amount of ability

1. A characteristic of low importance, a capacity possessed by two thirds of the population, a low degree of ability

0. The characteristic not required, or required only to a very slight degree

FIG. 2-2. A typical schedule used in estimated worker analysis.

Criteria for the Evaluation of Techniques of Estimated Worker Analysis. Procedures of estimated worker analysis differ greatly in the precision with which they evaluate worker characteristics. It is therefore necessary

to have standards by which any technique can be assessed. Three important standards are the reliability of the ratings of importance of the characteristics, the extent to which ratings of different worker characteristics are independent, and the extent to which the ratings are valid.

Reliability of Ratings of Importance. In worker analysis, reliability would be indicated by the extent to which different analysts agree concerning the importance of the various characteristics. Assuming that the analysts are of high caliber and that specific, well-defined jobs are being studied, the degree of agreement among analysts is an important means of evaluating a particular technique. Perfect agreement in all worker characteristics, even among highly selected analysts, certainly cannot be expected. The usual procedure in cases of disagreement is to invite open discussion among the analysts. By this means the causes of disagreement can be discovered and resolved, and deficiencies in the procedures used in the worker analysis can be pointed up.

Unreliability can usually be traced to one or more of the following: the nebulous character of the job, the lack of understanding of the worker characteristic being considered, or the relative importance of the characteristic as judged by the analyst. To be sure, disagreement may be due to lack of ability or training on the part of the analysts, but incapable or untrained personnel should not be utilized.

The reliability with which worker requirements are rated varies with the type of job. For some jobs the ratings assigned by different analysts agree quite closely while for others there is sharp disagreement. If the

Table 2-3. Distributions of Correlations between Ratings of Different Analysts on 38 Worker Characteristics for Five Jobs

Correlations between ratings by different analysts	Oven man	Typist	Crane operator	Power shovel operator	Waitress
.90 to 1.00	35	21	15	14	8
.80 to .89		5	6		7
.70 to .79	1	1	4		5
.60 to .69		3	2	5	3
.50 to .59					2
.40 to .49		2	1	7	3
.30 to .39		4	5		
.20 to .29			2	1	2
.10 to .19			2	8	2
.00 to .09	2	2	1	3	6

analysts are trained and experienced, this disagreement probably means that the jobs are not clearly defined. In Table 2-3 are shown the correlations between ratings made by different analysts who were highly experienced and were well trained. For the job of oven man these analysts showed very high agreement on 36 out of 38 characteristics rated. On the other hand, for the job of waitress the analysts could agree on only a few characteristics. Agreement for the other jobs was in between these two extremes. These results indicate that comparisons between requirements of different jobs must be made with caution since some jobs are more accurately described than others.

The reliability with which worker requirements are rated also varies with the kind of characteristic being judged. Illustrative results for experienced analysts are shown in Table 2-4. Here it will be seen, for example, that numerical facility is more reliably rated than eye-hand coordination. More definite statements, then, can be made about the importance of the former than about the importance of the latter.

Table 2-4. Distributions of Correlations between Ratings by Different Analysts on Five Worker Characteristics for 35 Jobs

Correlations between ratings by different analysts	Numerical facility	Memory for directions	Finger dexterity	Bimanual coordination	Eye-hand coordination
.90 to 1.00	27	22	19	11	7
.80 to .89	3	4	3	3	2
.70 to .79	1	1	2	2	1
.60 to .69		3	1	1	2
.50 to .59		1	1	1	1
.40 to .49		2	1		
.30 to .39	1		2	1	2
.20 to .29	1		2	3	2
.10 to .19				4	6
.00 to .09	2	2	4	9	12

Reliability is also affected by the relative importance of the characteristic as perceived by the analyst. The data in Table 2-5 represent the findings from worker analyses performed on three different industrial jobs. They show the relationship between the degree to which analysts agreed in their ratings of worker characteristics and the average importance they assigned to these characteristics. It will be observed that the greatest agreement occurs with characteristics that are considered

Table 2-5. Relationship between Estimated Degree of Importance of 41 Worker Characteristics on Three Different Jobs and the Reliability of Estimation

Correlations between ratings by different analysts	0	1	2	3	4	5	6	7	8	9	10
.90 to 1.00	51	17									
.80 to .89	1	9									
.70 to .79		6									
.60 to .69		1	7						1		2
.50 to .59			1								
.40 to .49			2						1		
.30 to .39			1	2							
.20 to .29				2				1		1	
.10 to .19		1		3	1			1			
.00 to .09				1	5	1	1	3			
	Unimportant								Important		

Estimated degree of importance of worker characteristic

unimportant. With those considered to be of moderate importance agreement is poor, and for characteristics considered to be important agreement increases to a moderate degree. It can therefore be expected that the reliability with which worker characteristics are rated will be greatest for those characteristics that either are considered to be very unimportant or very important.

Independence of Ratings of Different Characteristics. There is little point in having worker analysts evaluate jobs in terms of a variety of characteristics if the analysts are unable to discriminate between these characteristics. Although it may be true that a particular job requires strength of hands but not strength of arms, the analyst undoubtedly would rate both types as being important if he is unable to judge reliably which muscle groups are being used. In such a case the analyst would check both arm and hand strength as being important in any job where either one or the other was used. For one job both characteristics would be given a rating of "no importance," for another job both would be given a rating of "some importance," and for a third job both would be given a rating of "paramount importance." If a number of jobs were analyzed, the coefficient of correlation between ratings on these two characteristics would be 1.00. Under such circumstances, ratings of importance on the single trait "strength of hands and arms" would give the same description as ratings on either strength of hands or strength of arms given separately.

Lack of discriminability is due to several factors. It may be due to inability of the analysts, through sheer human obtuseness, to differentiate traits that are truly different. In this event the remedy is a more careful selection and training of analysts. Even skilled analysts, however, will

have difficulty in differentiating between worker characteristics if they are poorly defined or if they represent traits that are too general in nature. The less exactly the traits are defined, the less basis there will be for the analysts to discriminate differences. The more generalized the traits are, the more meanings they will have in common, and as a consequence the more likely they are to be given the same ratings of importance for any particular job.

Jaspen has reported results of an investigation wherein he shows that the ratings on 20 traits are really reducible to six trait-groupings.[11] Worker analysts rated the importance of 20 characteristics for 275 skilled, semiskilled, and unskilled jobs. The correlations between ratings on the various characteristics were then studied. In Table 2-6 are given some

Table 2-6. Correlations between Certain Estimated Worker Characteristics

Characteristic	Strength				Manual dexterity			
	Strength of hands	Strength of arms	Strength of back	Strength of legs	Dexterity of fingers	Dexterity of hands	Eye-hand coord.	Independent movements
Strength:								
Strength of hands		.82	.50	.62	−.07	.44	.24	.25
Strength of arms			.88	.72	−.40	.23	.02	.12
Strength of back				.95	−.29	−.15	−.14	−.15
Strength of legs					−.21	−.08	−.06	−.16
Manual dexterity:								
Dexterity of fingers						.35	.62	.44
Dexterity of hands and arms							.81	.55
Eye-hand coordination								.67
Coordination of independent movements								

of the findings from Jaspen's investigation. It will be observed that a high degree of consistency exists among the characteristics of strength of hands, strength of arms, strength of back, and strength of legs. A job that was rated high on any one of these tended to be rated high on all others. A job rated low on one was assigned similar ratings on the others. These four characteristics form a group or cluster and all seem to be concerned with the same factor, namely, strength. A similar consistency exists among the other four characteristics of dexterity of fingers, dexterity of hands and arms, eye-hand coordination, and coordination of independent movements. These characteristics appear to have in common the factor of manual dexterity. It will also be observed from the table that the correlations between ratings of characteristics in the

strength group and those in the dexterity group are, on the average, quite low, and hence these two groups clearly are concerned with different factors.

By observing in this way consistencies among ratings of the 20 different characteristics rated, Jaspen was able to differentiate the following six different groups of worker characteristics: strength, intelligence, inspection, physically unpleasant working conditions, manual dexterity, and mechanical information. Had the jobs been rated on these six characteristics rather than on the original 20, about the same amount of meaningful information might have been obtained.

Validity of Ratings of Worker Characteristics. If the traits that in actual fact are important for successful performance of a job are not so noted, then the worker analysis obviously is inadequate. Similarly, if the degree of importance of any characteristic is either underestimated or overestimated, the analysis cannot be considered a satisfactory one. The problem here is essentially one of validity. Do the estimated requirements of the job correspond to the actual requirements?

From a practical point of view such an evaluation is most difficult to make, and in some cases may be impossible since there is no final authority on the demands of the job to which the investigators can appeal. Indeed, if there were, this final authority would be used rather than estimations made by worker analysts. For the most part, the best that can be done at the present time is to resort to indirect measures of the validity of the analysts' judgments. Since these measures will be indirect, they should be interpreted with considerable caution.

In a worker analysis made for the purpose of developing selection tests, for example, the importance of the various traits as estimated by the analysts can be compared with the coefficients of correlation between scores on the tests used to measure these traits and success on the job, *i.e.*, the validity coefficients of the tests. If tests for those characteristics considered very important for the job have at least moderately high validity coefficients, and those for characteristics considered only fairly important have lower validity coefficients, then the correspondence between the estimated requirements of the job and the validity coefficients of the tests strongly suggests that the results of the worker analysis were satisfactory.

For purposes of illustration, Table 2-7 shows a comparison of the relative importance of several characteristics deemed necessary for the job of inspector-packer and the validity coefficients of tests selected as probably good measures of these characteristics. For the five traits considered to be of some importance there appears to be a rough relationship between estimated degree of importance and the relative size of the validity coefficients of the tests employed to measure them. In addition, for five

Table 2-7. The Importance of Worker Characteristics for the Job of Inspector-Packer as Judged by a Worker Analyst Compared to the Validity Coefficient of Tests Selected to Measure These Characteristics

Worker characteristic	Rank in importance of characteristic as estimated by analyst	Validity coefficient: Rank	Validity coefficient: r	Test selected to measure the characteristic
Finger dexterity	1	2	.50	Peg board
Perception of differences in details	2	3	.42	Name and number comparison
Perception of size and form	3	1	.57	Spatial relations
Eye-hand coordination	4	4	.40	Tapping and dotting
Dexterity of hands and arms	5	5	.24	Block placing
Average validity coefficient of 5 additional tests measuring traits considered to be of negligible importance			.13	

tests measuring traits believed to be of little importance for the job, the average validity coefficient is quite low. The results of this particular investigation suggest that the technique of worker analysis employed yielded satisfactory estimates of the degree of importance of the various traits considered.

If either the tests used to measure the traits or the measures of proficiency on the job are in fact not valid indices, the evaluation cannot be considered a satisfactory one. For example, suppose manual dexterity of the sort involved in well-coordinated movements of the wrist is important for the job and is so noted by the analyst. If the test designed to measure manual dexterity involves rapid movements of the hands but little coordination of wrist movements, a low correlation between test scores and success on the job would not be surprising. Similarly, if success on such a job was measured in terms of superiors' ratings, and these ratings did not take production records into sufficient account, even an adequate test of the trait would show little validity.

Sometimes certain facts like labor turnover and accident rate furnish information on the validity of worker analyses. Certain analyses are carried out primarily to get information concerning the physical demands of various jobs.[9] The aim of such analyses is to provide information about the jobs that will be useful in the placement of physically handicapped persons. Knowledge of the physical demands of jobs on the one hand,

and an individual's physical capacities on the other, makes possible a more effective job placement of the individual. If the analytic procedure gives accurate and complete information, then a decrease in sickness rate, accident rate, or labor turnover should be expected from the improved placement of the workers.[10]

PROBLEMS AND TECHNIQUES IN THE STUDY OF MEASURED WORKER CHARACTERISTICS

Objective Test Technique. More than thirty years ago Link suggested that the importance of worker characteristics might be indicated by objective measurements through psychological tests rather than through subjective estimations.[12] By testing workers on a given job with standard psychological tests information could be obtained which would give indications of the relative importance of various abilities and traits. A job profile could be constructed from actual objective measurements, rather than from subjective ratings.

Attractive as this procedure appears, few systematic analyses utilizing it have been performed. The approach of measured worker characteristics is extremely expensive since a variety of tests requiring some five or more hours of testing time would have to be administered to a fairly large number of workers. The administrative problems would be substantial. Furthermore, as later discussion will show, it is not altogether clear just what types of measurements should be utilized. Finally, there are a number of methodological problems that must be solved. In the face of these difficulties it is not surprising that the simpler approach of estimated worker characteristics has proved more attractive.

The Problem of Sampling Worker Abilities and Traits. In the approach of measured characteristics it is important to utilize a wide variety of tests so that all manner of abilities can be evaluated in terms of their importance. If the basic abilities and personality traits of man were known the task would be simplified. Just those tests known to measure these abilities and traits would need to be used. Unfortunately the basic abilities and personality traits of man, if such exist, are not known. The evidence from statistical and experimental studies of individual differences suggests that man possesses a number of separate and independent abilities and traits, but knowledge concerning how many there are or what their nature is is lacking at the present time. As a consequence, when worker requirements are described by means of measurements obtained through psychological tests it is necessary that the tests employed cover a wide variety of abilities and traits so that the possibility of excluding important ones is minimized.

Average-score Technique. The most extensive attempt to describe worker qualifications through measured characteristics has been that of the Employment Stabilization Research Institute of the University of Minnesota, using the average-score technique.[8] The investigators in this institute standardized a number of psychological tests by administering them to a representative cross section of the general employed population. The same tests were then administered to workers on a number of different jobs, and the average scores made on the different jobs were

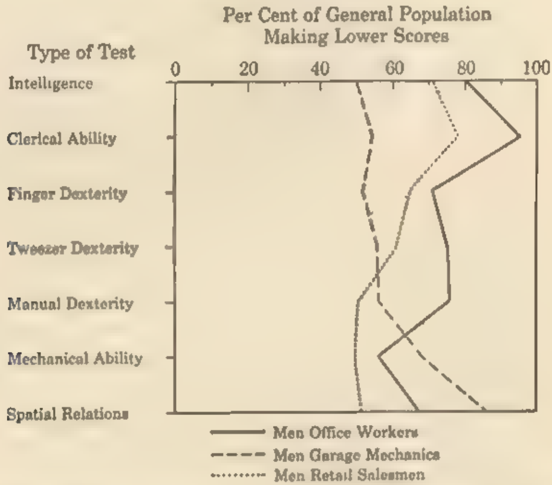


FIG. 2-3. Average scores on several psychological tests made by different occupational groups.

calculated. Typical results obtained from three groups of workers are shown in Fig. 2-3. The average score on each test is taken as the index of importance of the abilities or traits measured by the test. Thus from the data shown in Fig. 2-3 it would be said that for garage mechanics mechanical and spatial abilities are of considerable importance while intelligence, clerical ability, and finger dexterity are less important.

Paterson has examined the basic assumptions underlying the average-score technique.[14] Among other things he has pointed out that the assumption must be made that tests on which workers on a given job are superior should also show a positive correlation with success on that job. If this were not true, then an occupational pattern of test scores would be meaningless in terms of ability associated with successful work on that job. The high scores earned by a particular job group would have to be explained on the basis of other factors. It should be expected then that the higher the scores earned on a test by a group the higher will be the validity of the test.

Dvorak has offered certain evidence in support of the validity of Paterson's assumption.[9] With several jobs she has shown that workers varying in success on the same job show similar patterns of test scores, but that the poorer workers uniformly earn lower scores than the better ones. There are other instances, however, where workers on different jobs make widely different scores on the same tests, but the scores, nevertheless, show little or no relationship with success on the job. Some examples of this are shown in Table 2-8.[10] At present, it cannot be concluded that

Table 2-8. Differentiation of Workers on Various Jobs by Tests Which Show an Insignificant Correlation with Job Success

Test and job	Validity coefficient	Mean score	Standard deviation
O'Rourke Vocabulary Test:			
Card-punch-machine operator	.04	69	8
Bookkeeping-machine operator	−.07	59	16
Hand transcriber	.01	55	18
Department-store salesperson	.01	48	18
Cafeteria floor girl	.04	42	10
Power-sewing-machine operator	−.07	25	14
Minnesota Clerical Test—Number Comparison:			
Bookkeeping-machine operator	−.09	154	21
Calculating-machine operator	.10	140	16
Inspector-wrapper	.02	110	21
Cafeteria floor girl	.00	102	21
Hand transcriber	.08	95	26

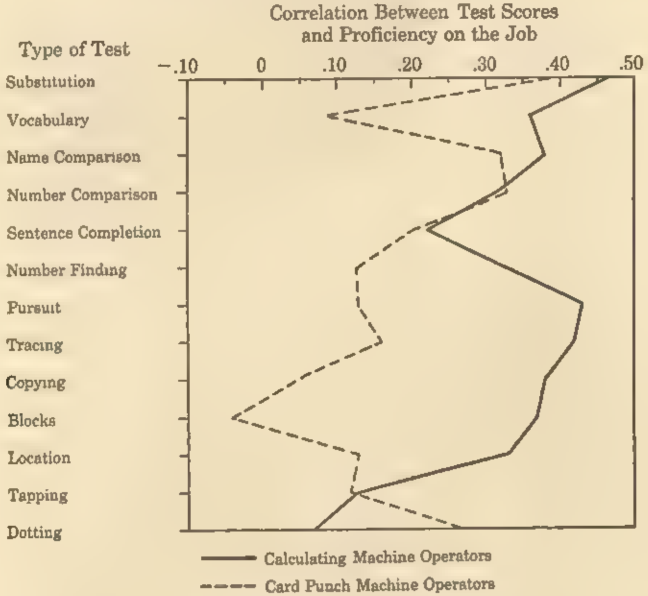
the assumption pointed out by Paterson has been demonstrated to have any general validity.

Among other factors that may result in workers on a given job having high scores on certain tests is differential sampling. Individuals with certain patterns of abilities and traits may seek work on a given job, rather than other jobs, not because their abilities are particularly suited to it, but because of social pressure, custom, convenience, availability, etc.

Test-validity Technique. In the discussion of the average-score technique it was implied that the most direct attack on the problem of measured worker characteristics would be to describe the importance of various tests in terms of the validity coefficients of these tests. While this approach has not been employed in any systematic way it is occa-

sionally utilized and appears to be very promising. By way of illustration, data have been drawn from a study wherein the same set of tests was administered to workers on two different jobs and the correlations between test scores and indices of success on the job were determined.[16] Job profiles, as shown in Table 2-9, could be constructed and conclusions drawn concerning the relative importance of various characteristics.*

Table 2-9. Coefficients of Correlation between Scores on Several Tests and Job Proficiency on Two Different Jobs



The test-validity technique of judging the importance of worker characteristics is subject to several important difficulties, some of which are methodological, and some conceptual. In many characteristics a given group of workers is likely to show a restricted range of talent when compared to the general population. For example, if only college graduates were hired for a clerical job, a test of arithmetic ability might show no relationship with success on the job because all individuals would possess a high degree of facility with numbers. On the other hand, in terms of finger dexterity the group might have about the same distribution of ability

* In order to express the importance of various abilities and traits in more fundamental terms the intercorrelations among tests and criteria could be subjected to a factor analysis. The loadings of the factors in the criterion would be the description of importance of the various primary characteristics. Several investigators have made such analyses with quite suggestive results.[18]

as exists in the general population. Supposing that both arithmetic ability and finger dexterity were equally important for job performance, the correlation between scores on the arithmetic test and job proficiency might be much lower than the correlation between finger-dexterity scores and job proficiency. This would be due to the fact that, in the college group, there is a much smaller range of talent in arithmetic ability than in finger dexterity.

A second difficulty arises out of the fact that job proficiency is never measured with perfect reliability. Differences in proficiency are much more reliably measured on some jobs than on others. Thus speed of typing ordinarily is quite a stable measure, while accidents on a job are unstable. For the job of typing, a test of reaction time may show no correlation with success because the ability is not important. For the job of driving a bus no correlation with success may again be found, not because reaction time is unimportant, but because proficiency on this job is unreliably determined.

A conceptual difficulty arises when scores on two different tests show the same degree of relationship with success on the job and yet the workers earn much higher scores on the one than on the other. The question can be raised as to whether the two abilities really are of equal importance. The logic of this situation has not been thought out and further clarification of the problem is in order.

Finally, it should be pointed out that the detection by tests of all abilities and traits required on a job seems improbable of achievement. Even though highly reliable indices of job proficiency are available, very high correlations are never obtained between scores on tests, used singly or in combination, and measures of job proficiency. Suppose that a large number of tests apparently covering all possible abilities and traits were administered to a group of workers on a given job, and the best combination of scores then obtained. It is very unlikely that the correlation between the scores and success on the job would exceed a value of .70 or .80. This means that because of imperfect knowledge and skill on the part of the analyst, it is impossible to measure all of the qualities that are important in the job performance. The test-validity technique is therefore defective in that it does not get complete information concerning the importance of worker characteristics.

Conclusions Concerning Measured Worker Characteristics. Although highly suggestive as a method of worker analysis, the study of measured worker characteristics is in its infancy. So little work has been done with the method that many of the basic problems have not been defined and many of the concepts have yet to be formulated. As further knowledge is gained concerning the nature of human abilities, the problems will become clearer and more exact methods will be developed. Primitive as

present methods are, however, the results appear to be superior to those obtained from the approach of estimated worker characteristics.

OCCUPATIONAL FAMILIES

Although for many years jobs have been grouped into various kinds of occupational families, it can be said that the study of the relationships among jobs is still in the early stages of development. With few exceptions problems in this area have been dealt with in a casual and unsystematic manner. Concepts have not been well formulated nor basic questions clearly stated.

A number of systems have been devised for the grouping together of jobs into families, but by and large the basis on which the grouping is accomplished is intuitive rather than empirical. A typical classification of this sort is as follows:

Professional and managerial	Sales
Semiprofessional, minor business, and supervisory	Skilled
	Semiskilled
Clerical	Unskilled

It is said that such a system is a logical one, but a close inspection of the assignment of jobs to these categories will reveal that the bases on which the groupings are made are quite nebulous. Thus it is said that "obviously" the jobs of carpenter, loom operator, and butcher require much more skill than the jobs of truck driver, electric-motor assembler, and service-station operator. Hence the former should be grouped together into the skilled trades and the latter into the semiskilled occupations. Obvious though this may be, no objective evidence is offered to show that in quantitative terms the jobs within one of the groups are more like each other than they are like the jobs in the other group.

Purposes for Which Jobs Are Grouped into Occupations. A systematic attack on the problem of occupational grouping can be initiated with a consideration of the principal purposes served by such grouping. The first and most obvious purpose for the grouping of jobs is for problems connected with personnel placement. With knowledge of which jobs are similar, more effective transfers of workers can be made both horizontally and vertically. An example where such knowledge is extremely useful is in the transfer of persons between civilian and military jobs. Information relative to similarities among jobs is also necessary for the evaluation of the past experience of applicants. Similarly, training programs can be more intelligently formulated when the similarities and differences among the jobs involved are understood.

The second major purpose for having well-defined occupational families is in connection with administrative matters. In many instances legisla-

tion is designed for certain classes of jobs but not for others. In the establishment of wage rates it is frequently desirable or even necessary to set up equal rates for similar kinds of jobs. In some instances such classifications are necessary as a basis for determining the jurisdiction of labor unions.

Types of Characteristics Used in Determining Similarity among Jobs. Jobs can be grouped into families on the basis of their similarity in any one of three general classes of characteristics. These are:

1. Administrative characteristics, such as hours of work, wages, responsibilities, etc.
2. Characteristics which the worker acquires in the performance of a job, such as the specific knowledges and skills that he learns
3. Characteristics which the worker possesses apart from those acquired from job experience, such as aptitudes, personality traits, education, etc.

On first thought it might be supposed that grouping jobs in terms of administrative characteristics would pose no great problems. This supposition, however, is not borne out in fact. For example, two jobs may both require a 40-hr. week, but one requires five 8-hr. days while the other requires five 7.4-hr. days and one 3-hr. day. Although the over-all hours are the same per week, their distributions throughout the week differ significantly in the demands placed upon the workers.

Much attention has been given to grouping jobs by listing common duties and responsibilities. A given duty or responsibility, however, may be much less important or consume much less time on one job than on another. Both an executive and his secretary use a letter file. A simple statement that this duty either is or is not performed, therefore, is not sufficient. A similar situation exists in describing responsibilities. Both the president of a concern and a foreman exercise supervision over men, but the supervision in the two cases is not of the same kind or of the same order.

The use of communality in knowledges and skills as a basis for grouping jobs together appears to have considerable practical value. Jobs so grouped could be considered equivalent with respect to such factors as difficulty in learning and amount of training time required. Presumably for those jobs involving essentially the same skills and knowledges workers could be interchanged readily and with little or no retraining. An individual who achieved proficiency on one job could be moved to another without much loss.

The grouping of jobs together in terms of communality in skills and knowledges, however, has three limitations. First of all it would be useful only when dealing with individuals who have had some previous work experience, an experience sufficiently great to bring them to fairly high levels of knowledge and skill. Secondly, merely because an individual once has achieved such levels it does not follow that he will retain them.

Unless a particular set of knowledges and skills are maintained through use, forgetting will lessen their value. Therefore, it does not follow that an individual can be transferred from one to another of two jobs closely related in terms of required knowledges and skills just because at one time he had achieved considerable proficiency on one of them. Thirdly, in using this basis for grouping jobs the implication is made that the amount of transfer from one job to another is large. While transfer of training is a clearly recognized psychological phenomenon, the amount of transfer from one job situation to another is by no means as great as is generally believed. Unless the two job situations are closely parallel, the extent of transfer will be relatively small.

Certainly one of the most fundamental ways of grouping jobs together is in terms of similarity of abilities and traits important in their performance. Basing the formation of occupational families upon communality in abilities and traits has a great deal of meaning. Thus an individual who has successfully performed one job in a particular family can be expected to possess the potentiality for successful performance in any other member. Aptitude tests that have been found to be valid for the prediction of success in one job presumably will have substantial validity for others in the same family. According to this method it is not expected that a worker can perform adequately immediately upon being transferred to a related job. Rather, it is maintained that he has the necessary aptitudes to learn to perform adequately.

Job Families Based on Estimated Worker Characteristics. Several attempts have been made to group jobs together in terms of similarity in estimated worker characteristics. While the task would seem to be a simple one, merely involving the grouping together of those jobs which have similar profiles, it is apparent from the studies thus far made that clear-cut groups do not emerge. By way of example a study by Coombs and Satter may be cited.[4] These investigators analyzed the similarities and differences among 20 office occupations rated on 18 worker characteristics. The degree of similarity between each pair of jobs was expressed in terms of the correlation between the ratings of the worker characteristics. A high correlation signified very similar job profiles, and a low correlation dissimilar profiles.

An analysis of the results indicated that the jobs were quite similar in terms of estimated requirements, the average of the correlations among the jobs being about .60. This high degree of similarity might simply be due to the analyst's inability to distinguish differences, or it might mean that the requirements of the various jobs actually are quite similar. Over and above this general similarity there was some tendency for the jobs to form into four clusters which could be defined as self-responsible jobs, routine jobs, skilled machine-operation jobs, and clerical jobs. However,

in several cases jobs were found to be about equally similar to two clusters, and in others jobs were found to be quite dissimilar.

Job Families Based on Measured Worker Characteristics. Just as jobs can be grouped together in terms of estimated worker characteristics so they can be grouped with respect to similarity in measured characteristics. There have been few attempts to use this method. A very suggestive pioneering study in this area is reported by Thorndike and Norris.[18] These investigators analyzed the results from a battery of tests administered to large numbers of Air Force personnel in 30 military occupational specialties. Knowing the correlations between the scores on the various tests and the indices of success on the various jobs it was possible to group together those jobs that require similar abilities and traits. As in the study by Coombs and Satter, similarity between jobs was described in terms of the correlation between job requirements.

There appeared to be some general similarity among the various jobs, the average of the correlations between jobs being about .35. An analysis of the correlations among the jobs revealed that they tended to fall into five clusters. For administrative and training purposes in the Air Force the 30 jobs had been grouped into seven categories as follows: clerical, craftsman, electronics technician, equipment operator, mechanical, radio operator, and technician specialty. In Table 2-10 the five clusters of jobs which emerged from the analysis are compared with the seven administrative categories. It will be noted that there is little correspondence between the two types of categories. It can therefore be concluded that the organization of jobs into categories for administrative purposes showed little relationship with the organization of these jobs based on measured worker characteristics.

The Dictionary of Occupational Titles. No discussion of occupational families would be complete without reference to the "Dictionary of Occupational Titles" developed by the United States Employment Service.[19] Some fifteen years ago this organization initiated a wide-scaled study of the nature and requirements of jobs throughout the country. Thousands of jobs ranging from professional to unskilled were analyzed and described. This stupendous undertaking has been a continuing function of the U.S. Department of Labor, and the result has been an increasingly greater understanding of jobs and their interrelationships.

Since one purpose of this program was to provide information which would be of aid in placing persons, the attempt was made to group together those jobs among which workers would transfer most readily. The Employment Service made the reasonable assumption that in so far as the interchangeability of jobs is concerned the job specifications constitute the salient features for some jobs whereas the worker specifications

Table 2-10. Relationship between Job Categories Based upon Administrative Practices and Job Clusters Based on Similarities in Measured Worker Characteristics

Job categories for administrative purposes	Job clusters based on similarity in measured worker characteristics				
	A	B	C	D	E
Clerical	Clerk-typist Stenographer				
Craftsman		Fabric and dope mechanic		Welder Airplane sheet-metal worker Carpenter Electrician Machinist Plumber Sheet-metal worker	Parachute rigger Radar mechanic
Electronics technician	Radio mechanic				Remote-control-turret mechanic
Equipment operator		Construction-machine operator			Fire fighter & crash rescueman
Mechanical	Airplane & engine mechanic 74700		Diesel mechanic Engineman		Airplane electrical mechanic Airplane & engine mechanic 74710 Jet airplane & engine mechanic Airplane instrument mechanic Airplane propeller mechanic Automotive-equipment technician
Radio operator	Control-tower operator AACS Radio operator HMS				
Technician specialty	Weather observer				Photographer

are the important consideration for other jobs. Although no single criterion was used in grouping jobs into occupations, the emphasis was on worker specifications. On the whole, then, jobs with the same worker requirements were classified together. However, duties, types of machines used, nature of the materials worked with, nature of the articles produced, and working conditions also were taken into consideration.

Table 2-11. General Outline of Classification of Jobs in the U.S. Employment Service "Dictionary of Occupational Titles"

0- Professional and managerial occupations
  0-0 through 0-3 Professional occupations
  0-4 through 0-6 Semiprofessional occupations
  0-7 through 0-9 Managerial and official occupations

1- Clerical and sales occupations
  1-0 through 1-4 Clerical and kindred occupations
  1-5 through 1-9 Sales and kindred occupations

2- Service occupations
  2-0 Domestic-service occupations
  2-2 through 2-5 Personal-service occupations
  2-6 Protective-service occupations
  2-8 through 2-9 Building-service workers and porters

3- Agriculture, fishery, forestry, and kindred occupations
  3-0 through 3-4 Agricultural, horticultural, and kindred occupations
  3-8 Fishery
  3-9 Forestry, hunting, and trapping

(4- and 5- Skilled; 6- and 7- Semiskilled; 8- and 9- Unskilled)

4-, 6-, 8-00. through 5-, 7-, 9-18. Occupations in manufacturing and related activities
  4-, 6-, 9-01. through -10. Production of food products
  4-, 6-, 8-12. Manufacture of tobacco
  4-, 6-, 8-14. through -10. Manufacture of textiles
  4-, 6-, 8-21. through -27. Production of fabricated textile products
  4-, 6-, 8-29. through -39. Production of lumber and lumber products
  4-, 6-, 8-41. through -42. Production of paper and paper goods
  4-, 6-, 8-44. through -49. Printing
  4-, 6-, 8-50. through -53. Production of chemicals and chemical products
  4-, 6-, 8-56. through -56. Production of petroleum and coal products
  4-, 6-, 8-57. Production of rubber goods
  4-, 6-, 8-59. through -62. Manufacture of leather and leather products
  4-, 6-, 8-65. through -70. Production of stone, clay, and glass products
  4-, 6-, 8-71. through -95. Metalworking occupations
  4-, 6-, 8-97 through -99.
    and 5-, 7-, 9-00. Manufacture of electrical equipment
  5-, 7-, 9-02. through -05. Manufacture of transportation equipment
  5-, 7-, 9-08. through -13. Manufacture of miscellaneous products
  5-, 7-, 9-16. through -18. Miscellaneous manufacturing occupations

5-, 7-, 9-20. through 5-, 7-, 9-61. Occupations in nonmanufacturing occupations
  5-, 7-, 9-20. through -22. Extraction of minerals
  5-, 7-, 9-33. Construction occupations
  5-, 7-, 9-35. through -49. Transportation occupations
  5-, 7-, 9-51. through -54. Communication and utility occupations
  5-, 7-, 9-55. through -60. Trade and service occupations
  5-, 7-, 9-61. Public-service occupations

5-, 7-, 9-63 through 5-, 7-, 9-89. Miscellaneous occupations

5-91 through 5-99. Foremen

7-93 through 7-99. Apprentices

Some forty thousand jobs are defined in the "Dictionary of Occupational Titles," and each is assigned either a five- or a six-digit code number. The jobs are classified by means of their code numbers into seven major occupational groups, each of which, in turn, is divided and subdivided into smaller groupings with increasingly greater similarity. The code numbers run from 0-01.10 for the job of cost accountant to 9-98.91 for gravedigger. The first digit in the code is the most general in meaning, referring to one of the seven major occupational groups. The subsequent digits refer, in turn, to finer and finer classifications.

Table 2-11 presents a general outline of the system of classification of occupations. It will be noted that for jobs with code numbers beginning with 4 through 9, the classification is twofold; in one grouping, the jobs are classified according to skill, and in the other, according to industry or type of activity. There is no hierarchical order of jobs from "good" to "poor," or from "desirable" to "undesirable." The intent was to base the classification solely on the similarity and interchangeability of the jobs.

The system of occupational families as given in the "Dictionary of Occupational Titles" is very serviceable because of its complete coverage of jobs. It has a further advantage in that each of the jobs is defined and a terminology of job titles is provided. There are, however, some shortcomings. Since estimated worker characteristics were employed, the extent of error due to personal bias and judgment is unknown. Many experienced personnel workers disagree with the groupings of certain of the jobs as given in the "Dictionary." This suggests either errors of judgment or incomplete job information. Nevertheless, the "Dictionary of Occupational Titles" is a most helpful document in personnel placement, and undoubtedly will continue for many years to be the standard instrument for describing similarities and interrelationships among jobs.

REFERENCES

1. Charters, W. W., and I. B. Whitley: "Analysis of Secretarial Duties and Traits," Williams & Wilkins, 1924.
2. Christensen, J. M.: A sampling technique for use in activity analysis, *Personnel Psychol.*, **3**, 361–368, 1950.
3. Colby, L. B.: Milk distributors seek to make 30,000 route robots into creative salesmen, *Sales Mgmt.*, June, 1937.
4. Coombs, C. H., and G. A. Satter: A factorial approach to job families, *Psychometrica*, **14**, 33–42, 1949.
5. Division of Occupational Analysis and Manning Tables: "Training and Reference Manual for Job Analysis," War Manpower Commission, Washington, D.C., 1944.
6. Dvorak, B. J.: Differential occupational ability patterns, *Univ. Minn. Bull. Employ. Stab. Research Inst.*, **3**, No. 8, 1935.

7. Flanagan, J. C.: Critical requirements: a new approach to employee evaluation, *Personnel Psychol.*, **2**, 419–425, 1949.
8. Gilbreth, F. B.: "Motion Study," Van Nostrand, 1911.
9. Hanman, B.: "Physical Capacities and Job Placement," Nordisk Rotogravyr, Stockholm, 1951.
10. Hanman, B.: Matching the physical characteristics of workers and jobs, *Ind. Med.*, **14**, 405–430, 1945.
11. Jaspen, N.: A factor study of worker characteristics, *J. Appl. Psychol.*, **33**, 449–459, 1949.
12. Link, H. C.: "Employment Psychology," Macmillan, 1920.
13. Michael, W. B.: Factor analyses of tests and criteria, *Psychol. Monograph*, **63**, No. 298, 1949.
14. Paterson, D. G.: Scouting along the frontier, *Occupations*, **12**, 30–33, 1934.
15. Shartle, C. L.: "Occupational Information," 2d ed., Prentice-Hall, 1952.
16. Stead, W. H., and C. L. Shartle: "Occupational Counseling Techniques," American Book, 1940.
17. Taylor, F. W.: "Common Sense Applied to Motion and Time Study," Harper, 1911.
18. Thorndike, R. L., and R. C. Norris: "Empirical Evidence on Air Force Career Fields," Air Training Command, Human Resources Research Center, Research Bull. 52–13, 1952.
19. U.S. Department of Labor: "Dictionary of Occupational Titles," 1939 (and later editions).
20. Utter, R. F.: Relation of personality and character requirements to jobs in a civil service agency, *J. Appl. Psychol.*, **31**, 651–654, 1947.
21. Viteles, M. S.: Job specifications and diagnostic tests of job competency designed for the auditing division of a street railway company, *Psychol. Clinic*, **14**, 83–105, 1923.

CHAPTER 3

The Measurement of Job Proficiency

Having a description of the job in terms of the job and worker specifications, the next step is to devise and put into operation procedures for measuring the degree to which workers are successful in executing the duties and responsibilities of the job. What is desired are indices or yardsticks by means of which the relative success of workers can be gauged. Whenever possible, descriptions of proficiency should be given in quantitative terms in order that accurate comparisons can be made between individuals, and in order that each individual's performance can be evaluated against common standards of performance.

Before problems involved in the measurement of job proficiency and methods for achieving such measurements are discussed, it will be well to consider the fundamental question of whether it is either necessary or desirable to accomplish such measurements. Without thinking through the situation the answer to this question will generally be in the affirmative. On further reflection, however, a negative answer might appear to be more reasonable. If measurements are to be made they should be made with some specific purpose in mind. But many times it is found that workers are differentiated in terms of their proficiency with no reference to any specific purpose. The records are simply filed away for "future reference." The purpose for which measurements of job proficiency are made will be a factor in determining the kind of index used. This purpose will be connected with some contemplated administrative action or planning, and an index devised to serve one purpose might be valueless for another. An index of job proficiency devised with no particular purpose in mind is likely to serve none adequately.

A further justification for not measuring the proficiency of workers arises out of those situations where only inadequate measurements are obtainable. Inadequate measurements mean inaccurate quantitative descriptions of proficiency. If administrative actions are taken on the basis of such measurements, they might achieve results at variance with those desired. In many instances it might be better to have no information at all than to have faulty information.

The discussion in this chapter is predicated upon an affirmative answer to the question of whether it is necessary or desirable to measure the proficiency of workers on their jobs.

PURPOSES SERVED BY INDICES OF JOB PROFICIENCY

Proficiency measurements serve many useful purposes. Following is a list of purposes for which indices of job proficiency are used:

1. As a basis for the promotion, separation, or transfer of workers
2. As a means for evaluating the effectiveness of the work of an organization, or of its separate departments
3. As a basis for the payment of piecework wages, bonuses, etc.
4. As a means for evaluating the effectiveness of different methods of work, tools and equipment, and conditions of work
5. As a means for evaluating the effectiveness of devices used for the selection and classification of workers
6. As a basis for planning work schedules, estimating labor costs, etc.

In studying the foregoing list, the reader will note the fact that measures of job proficiency are needed throughout most of the phases of business and industry. This is not being stated as a new idea; it is an old idea in business and industry. Although many procedures have been devised for evaluating worker performance, the problem for the most part has not been given careful scientific scrutiny. Relatively few organizations utilize procedures for measuring job proficiency which have been systematically developed and which have known dependability. Sound methodological techniques are now available by which the dependability of any measure of proficiency can be determined, and an understanding of these techniques should go far toward effecting an improvement in present-day methods of measuring worker proficiency.

STANDARDS FOR EVALUATING INDICES OF JOB PROFICIENCY

Standards are needed for evaluating measures of proficiency. Quantitative records of worker performance can usually be obtained on most jobs. Once they are available there may be a tendency to consider the problem of job proficiency measurement as solved. Actually a very significant question still remains to be answered, namely, are the indices available really worthwhile. An index should meet certain standards before it is considered adequate.[2] The two most important criteria to be met are adequate validity and adequate reliability.

Validity. The validity of a measuring device refers to the degree to which the device measures what it is supposed to measure. The validity of an index of job proficiency refers to the degree to which the index

represents important characteristics of the job it purports to measure.[5] If the index involves items not contained in the actual work situation or items concerned with minor aspects of the job, the order of individual differences among the workers in *measured* job proficiency will bear little similarity to the order of individual differences of these workers in *actual* job proficiency. For example, suppose the proficiency of typists, who normally work in a noisy and crowded office, is taken as their performance on a work-sample test wherein they transcribe standard material for half an hour in a room by themselves under quiet conditions. This would not be considered a valid index of job performance since the test performance is obtained under non-representative conditions. In such a case the normal distracting influences, which operate differentially to reduce the output of some, but not all, workers on the job, would not be present in the test and hence would not contribute to the proficiency measurement.

Many indices of job proficiency are incomplete. For example, measures of output, even though obtained in the actual working situation, would be a deficient criterion if success on the job were mainly dependent upon the workers' ability to get along with their fellows. Before output could be considered a valid and representative index, it first would have to be demonstrated that any differences in success on the job attributable to ability to get along with other workers were adequately reflected in the differences obtained in the output criterion. Actually this factor of getting along with others is of considerable importance in job success. Causes of separation of workers from their jobs can be credited with reflecting factors concerned with worker success. In a large proportion of the cases the reason given for the separation of employees is not lack of ability to do the work but rather some defect or limitation of personality. Seldom, however, are attempts made to include this factor in the measurement of success on the job. If the measure of proficiency is to represent the job, it is of prime importance that all factors important in determining worker success be included.

In some cases an index of job proficiency might be contaminated by the operation of other conditioning factors. A supervisor's rating of a worker might be unduly high because he remembers that the worker made a high score on an employment test given three months earlier. The amount of a worker's production is influenced by the favorableness of the conditions under which he works and the quality of the tools and equipment which he is furnished. One bus driver may have a good safety record because his route covers suburban districts, while another driver would have a poor record because he operates through crowded traffic conditions. Indices of job proficiency contaminated by other factors will suffer in terms of validity.

The only practical way of judging representativeness is to check the measured characteristics against the important aspects of the job as revealed by the job specifications. Since some aspects of the job will be easier to measure than others, there will be a tendency to overlook certain less measurable but possibly very important aspects. In this connection it should be pointed out that a qualitative estimate of an important but difficult-to-measure aspect of a job may prove a better index than an accurate quantitative measure of a minor aspect of the job. A careful study of the job of department-store salesperson might reveal that the successful salesperson is not always the one who makes the greatest number or amount of sales. It may be one who makes such a favorable impression on the customer that, even though she does not purchase from that salesperson, she does purchase from another in some other department of that store rather than go to another store.

Representative measures of the less tangible and more qualitative factors contributing to job success sometimes can be obtained by using customer reactions.[8] For example, ratings or some other form of evaluation of a salesperson may be obtained from established customers who have been served by that salesperson on several occasions. Such a method is applicable to most retail merchandising situations. Similarly, instead of estimating the success of a foreman by obtaining a rating from his superior, a more valid measure might be obtained from the reactions of those working under him.

To achieve representativeness may require that different workers in the same department be evaluated by means of entirely different performance measures. Stead and Shartle have pointed out that successful measurement of proficiency sometimes requires more than one type of criterion within a department.[10] These authors have argued that success has many meanings. One worker may be considered successful because he raises the morale of the entire group, another because of his ability to do odd jobs in addition to his regular job, and a third because his production figures are always high.

Situations wherein several indices of proficiency might be required are very likely to arise in a job that is not rigidly defined, or in one where the employees have interchanged duties or do not work according to the specified duties of the job. The example of retail salespersons discussed in a previous paragraph serves as a case in point. One salesperson might be considered successful even though his sales are low, if he builds up such good will in the customers that they buy most of their goods in that particular store. The utilization of different criteria for evaluating the contribution of different workers within a given group has not received the systematic study that it merits.

Reliability. The reliability of a measure of proficiency is another important factor in determining its adequacy. By reliability is meant consistency or reproducibility. A measuring device is reliable if it gives the same quantitative description of an individual on two or more occasions. It is presumed, of course, that the characteristic being measured remains constant. The reliability of a measure should be determined by statistical analysis and not merely be assumed or, even worse, be totally ignored. It is often believed that a measure of proficiency will necessarily have high reliability because it is based on the worker's actual performance on the job. This is not necessarily true. Indeed, in many instances the reliability of the output of workers is so low that differences in performance of workers can largely be ascribed to chance factors.

The reliability of a measure of proficiency may be estimated either by repeating the measurement twice on the same group of employees and obtaining the coefficient of correlation between the two repetitions, or by estimating the reliability of the measure from correlating the scores obtained by dividing the proficiency records into comparable halves.[12] An example of the former method would be the correlation between the scores of a group of employees on two repetitions of the same work-sample test, or the correlation of amounts of sales for two periods of time. An example of the second method would be the corrected correlation between the output records of a group of employees on the odd and even months of a year, or the number of accidents of workers on the odd and even days of a month. Such coefficients of correlation are termed reliability coefficients.

Reliability and the Number of Measurements. Two important factors influence the size of the reliability coefficient, viz., the number of measurements obtained and the range of ability in the group being measured. In Table 3-1 may be seen the effects upon reliability of increasing the

Table 3-1. The Effect of Increasing the Number of Measurements upon the Reliability Coefficient

Original reliability coefficient	Number of times measure increased								
	2	3	4	5	6	7	8	9	10
.10	.18	.25	.31	.36	.40	.44	.47	.50	.53
.30	.46	.56	.63	.68	.72	.75	.77	.79	.81
.50	.67	.75	.80	.83	.86	.88	.89	.90	.91
.70	.82	.88	.90	.92	.93	.94	.95	.95	.96
.90	.95	.96	.97	.98	.98	.98	.99	.99	.99

number of independent measurements made on each worker being evaluated. It will be observed that, as the number of measurements increases, the reliability of the measuring device increases, but with diminishing returns. If, for example, the reliability coefficient of a 10-min. work-sample test were .50, increasing the length of the test to 1 hr. should increase the reliability to approximately .86. If the reliability of ratings by a single foreman were .70, the pooled ratings of four such foremen would probably be about .90.

It also will be noted from Table 3-1 that if a measure of proficiency has a low initial reliability, the number of additional measurements necessary to increase reliability to an adequate level will be too large to be practical. This is a serious deficiency when the proficiency measurements require considerable time or involve a large number of raters. Thus, if production records for a 6-month period had only a reliability of .30, it would mean that records would have to be taken on each worker for nearly five years in order to raise the reliability to .80. Similarly, if the reliability of the pooled ratings of three foremen were .30 (a not infrequent situation), 30 such foremen would have to be used to raise the reliability of the pooled ratings to as much as .80. Measures found to have a reliability so low as to make further increases in length impractical should be analyzed for correctable defects. Sometimes merely an improvement in the accuracy of record keeping or a more precise description of the performance will be sufficient to increase the initial reliability to a workable level.

Reliability and the Range of Ability. As the range of individual differences in ability or performance becomes smaller and smaller, the reliability coefficient decreases. Therefore, if a reliability coefficient of a measure of proficiency is determined from the performance of a group whose range of ability is less than that of the total group on which the measure is to be used, the reliability coefficient will be an underestimate. In order to obtain an adequate measure of reliability when all employees cannot be tested or records obtained from them, a sample of workers truly representing the range of ability of the total group should be used. No group restricted to any particular range of ability—superior, average, inferior—can supply the needed data.

The specific effects of reduction in range of individual differences upon the reliability coefficient may be seen in Table 3-2. Suppose that the production records of the upper 10 per cent and lower 10 per cent of a group of employees were not included in the determination of the reliability of output, then the range of individual differences would be 20 per cent less than that of the complete range. If the reliability coefficient of output, computed from this group with reduced range, were found to be .53, reference to Table 3-2 will show that the reliability coefficient for the total group would probably be .70. Table 3-2 also shows that the lower

Table 3-2. The Effects of Reduction in the Range of Talent or Individual Differences upon the Reliability Coefficient

Reliability coefficient of the total range	Reduced range expressed as a % of the original range				
	90%	80%	70%	60%	50%
.90	.88	.84	.80	.72	.60
.70	.63	.53	.39	.17	.00
.50	.38	.22	.00	.00	.00
.30	.14	.00	.00	.00	.00
.10	.00	.00	.00	.00	.00

the reliability coefficient, the more it will be decreased if computed on a reduced range.

The range of ability is an important factor in cases where output records are taken over a long period of time. The reliability of the criterion must be determined by using only those employees who remain on the job for the period under consideration. If there is any labor turnover, it probably will not be equal at all levels of job proficiency. It seems quite likely that those workers who leave the job will be either inferior workers who quit or are released because of low productivity, or superior workers who are promoted or leave the organization for better jobs. The range of ability of those workers who remain during the total time period, then, will be much smaller than the range of the whole department or plant personnel at any given time during that period.

KINDS OF PROFICIENCY MEASURES

A little reflection on worker performance will reveal that the ways in which proficiency can be measured are as varied as the ways in which the worker successfully contributes to his organization. Some indices have only a limited application because they involve factors that characterize only a few aspects of the job, whereas other indices have very wide usefulness because they involve a large proportion of the important job factors.

Many different types of devices have been developed for measuring the proficiency of workers. The indices that are most commonly utilized are amount and quality of production, work-sample tests, length of service, amount of training necessary, and ratings by superiors. With certain types

of jobs other specific indices are used. Thus accidents are an important index for bus drivers, number of complaints received for receptionists, and accuracy of determining course for navigators. In some instances the measurement is objective as is the case with amount of production where the number of units of work can be counted. In other instances the measurement is subjective as in ratings of neatness or leadership.

PRODUCTION AS A MEASURE OF PROFICIENCY

Amount and Quality of Output. Amount of output is probably the most commonly used measure of job proficiency and often is considered the best. Typical indices based on output are the number of units produced in a given interval of time, the time required to produce a given item, and amount of sales per given time period.

The amount of output should be considered in relation to the quality or accuracy of the work. An individual who turns out 100 units of work, all of which are perfect, must be considered better than one who, in the same time, turns out twice as many units, most of which are defective. A department-store salesperson who has a high sales record, but a large number of whose customers return purchased articles, cannot be considered as effective an employee as another whose sales are somewhat lower but whose returns are negligible. In all cases, then, the amount of output should be adjusted in terms of the quality of the work.

One method of making this adjustment involves an evaluation of the time, material, or cost involved in the imperfect or spoiled production. If an imperfect article must be sold at a reduction of 25 per cent from the market price of a perfect article, this reduction forms a quantitative basis for adjusting the output of workers turning out the imperfect articles. For each imperfect unit the worker would be credited with having completed three-quarters of a unit of work rather than a whole unit. If the work is of such a nature that any error committed must be corrected, *i.e.*, if the spoiled article must be discarded and another perfect one substituted in its place, then the time involved in producing the perfect article, or the cost of the materials in the discarded article, or both of these factors, can be used in adjusting the output.

Sometimes correcting an error requires more time than is involved in producing several correct articles, and in such a case the penalty of a mistake should be rather severe. The manner in which such corrections can be made may be illustrated by a study made by the U.S. Employment Service of Hollerith card-punch operators.[16] An investigation of errors indicated that on the average 13.75 cards could be punched while one error was being corrected. Consequently, in determining the output of a card-punch operator, 13.75 cards were subtracted from the total punched

for every card on which there was an error. For example, if an operator made two errors in punching 228 cards in 1 hr., 27.50 cards would be subtracted from the total, and the operator's output for the hour would be taken as 200.50.

The Problem of Comparable Units in Measures of Output. The basic prerequisite in the use of output as a measure of job success is that the units of work for all employees must be comparable.[17] By comparable is meant that the work units should be approximately equivalent in terms of the worker's time, accuracy, and energy. That is, if on job A each item that can be turned out with normal expenditure of worker energy requires 2 hr., then in an 8-hr. day the average worker should turn out four articles. Now if on job B with normal expenditure of worker energy the production of an item requires 30 min., then in an 8-hr. day the number of items completed by the average worker would be 16. On the average, then, one item of work on job A is equivalent to four items on job B. But the output of an individual in a bottling plant who is inspecting quart containers is not directly comparable with that of one who is inspecting pint containers, since the tasks may be of unequal difficulty in spite of the fact that they both involve the inspection of glass containers.

The simple index of number of items produced is not necessarily an adequate basis for comparing output. Even if the employees are working in the same department or handling similar items, the amount of work turned out is not a simple function of the worker's ability and effort. Differences in output among workers may be due to differences in the kind of work or working conditions, as well as to differences in the worker's ability or effort. In order to obtain an accurate measure of a worker's individual contribution in output the extraneous effects of differences in kind of work or working conditions should be eliminated or at least minimized.

A number of different methods have been developed for transmuting into comparable scales the output records of workers doing somewhat different tasks or working under different conditions. These methods all have in common the comparison of each individual's output with some kind of standard or quota. An individual's index of production is usually taken as the ratio of his production to the standard. Suppose, for example, that the standard set for an assembly operation is 20 units per hour. An individual who turns out 25 units per hour is assigned an *efficiency index* or *rating* of 125 (25/20 times 100). For another assembly the standard might be 25 units per hour and the individual who turns out 20 units per hour would be assigned an efficiency index or rating of 80 (20/25 times 100).

In many situations the standard is set in terms of the time to complete a given item rather than the number of items to be completed in a given

amount of time. Obviously, the one can be translated into the other. In certain cases different tasks on the same job may have different standards. For example, a power-sewing-machine operator at different times will work on women's cotton dresses, boys' overalls, and infants' playsuits. An efficiency rating for each worker is taken as the average of his efficiency ratings determined for each task separately. Workers are compared or evaluated by means of these so-called efficiency indices or ratings rather than in terms of the actual number of units of work they turn out.

It should be apparent that the degree to which production units on different jobs are made comparable will be a direct function of the goodness of the standards selected for the different jobs. A hypothetical example of the effects of inaccurate standards is given in Table 3-3. If

Table 3-3. Example of the Effects of Inaccurate Production Standards upon Efficiency Ratings

Standard	Job A performed by worker I	Job B performed by worker II
Inaccurate standard	4 per hour	10 per hour
Accurate standard	2 per hour	20 per hour
Actual production	3 per hour	15 per hour
% efficiency based on inaccurate standard	75	150
% efficiency based on accurate standard	150	75

the standard set for job A which is done by worker I is too liberal, and that set for job B which is done by worker II is too severe, then worker I's efficiency rating will be too high while that of worker II will be too low. Only when workers are doing the same tasks with equivalent machines, working conditions, and standards will their indices be perfectly comparable. Then, even though the standard is too high or too low, the error will be constant for all individuals. Under this set of circumstances, however, there is no need for a standard, as all factors outside of the worker are considered constant.

TYPES OF PRODUCTION STANDARDS

Standards of production are set in a variety of different ways. Sometimes the average output of the group of workers is used. In other cases

the output of specially selected individuals is used, to which a correction is applied so that the "average" is estimated. In some instances standards are set on the basis of experience gained from similar tasks. The industrial engineer's procedure of time study is perhaps the most commonly employed method. A relatively new attack on the problem is the use of what has been called experimentally determined times, wherein the time to make various kinds of motions is determined separately for each motion, and the standard for the total task is taken as the composite time for the component movements. Finally, standards are set by what has been termed rational analysis; the results of deliberations of "experts" or "interested parties."

With certain standards two factors are involved, viz., operation time and extra time allowances. This is particularly true with such procedures as time study. Operation time consists of the time required for the actual movements involved in completing the task, the speed of operation of machines, etc. Extra time is added to operation time in order to allow for unavoidable delays and work stoppages beyond the control of the worker, personal needs of the worker, and reduction in the individual's rate of work due to fatigue. It is to be noted that allowances made for these factors generally will be arbitrary. For example, in many cases there is no way of knowing the cause of a delay, and hence it is difficult to classify it as either avoidable or unavoidable. That such time allowances are arbitrarily determined is indicated by the variation found in them for the same jobs but in different organizations. Allowances for personal needs have been found to differ as much as 250 per cent, and for fatigue as much as 150 per cent.[14] In view of the difficulties inherent in the setting of extra time allowances they must be regarded with some reservations. Because they vary so much and because they are set arbitrarily, their use in setting standards can only mean that the standards themselves are necessarily inaccurate.

Average Production as a Standard. This method ordinarily is employed when the tasks performed by all workers are the same, or approximately the same, but where the units of measurement differ for different subgroups of workers, or where certain factors operate to influence differentially the output of different groups of workers. In a department store, for example, in one department selling articles for very small amounts, say 25 cents or less, the productivity of the salespersons might be measured in terms of the number of items sold per day. In another department selling expensive articles, productivity might be measured in terms of the dollar value of goods sold per day. Obviously, the two different units of productivity measurement are not directly comparable. It would be incorrect to say, therefore, that a salesperson in the first department who sold 700 items per day was either a better or poorer salesperson than one in the

second department who sold $500 worth of goods. Without some reference point to which each worker's productivity can be referred, such as the average number or amount of sales in the worker's department, comparisons would be impossible.

Similarly, in a factory where some workers operate older and slower machines and others newer and faster machines, direct comparisons in output between workers in the two groups would be fallacious even though the finished articles would be the same. The output of the first group would be expected to be less than that of the second group simply because their machines were slower. Some procedure is needed for adjusting the output of the workers in order to eliminate the error introduced by the differential effect on performance of the varying speeds of the machines.

One way in which adjustments in output are made for situations such as those described is to express each individual's output as a ratio of the average output of all the workers in his subgroup. Each worker's output is expressed as a percentage of the average for his department or subgroup, and variations between workers are measured in units of percentage. If it may be assumed that units of percentage based on average performance within a group are comparable for different groups, then the productivity of workers in different groups can be compared.

The manner in which these indices are computed may be illustrated by the following hypothetical example. Suppose that employees in a plant, all of whom are doing the same machine task, are assigned to one or another of three models of the machine. These models—A, B, and C—perform essentially the same operations but vary somewhat in their controls and speed of operation. Differences in amount of production may be expected from the workers in accordance with the characteristics of the machines assigned to them. In the left half of Fig. 3-1 are given the distributions of the production rates of workers assigned to each model of the machine. In general the production rate of workers assigned to machine A is lowest, and that of workers assigned to machine B is highest, the average rate of production being 220 for machine A, 300 for machine B, and 280 for machine C. The index of output for a person on machine A whose output is 242 would be 242/220 or 110. This would mean that this worker was better than average, since an average index would be 100. However, a person whose rate of production on machine B was 242 would have an index of output of 242/300 or 81, and clearly would be performing below the average.

If each worker's production is divided by the average production of all individuals working on the same type of machine, then a basis is obtained for learning about the relative productivity of workers using different models of machines. These computations have been performed

for the three groups of workers, and the distributions of indices of output are given in the right half of Fig. 3-1. It will be observed that this transformation renders the productivity of the three groups far more comparable than the simple production records of the workers.

There are certain conditions that must be fulfilled if this transformation procedure is to be satisfactory. The average production of each of the subgroups of workers must be based on a number of cases sufficiently large to be stable and reliable. With small groups of workers, as in the hypothetical case given above, the averages would be too unreliable to be

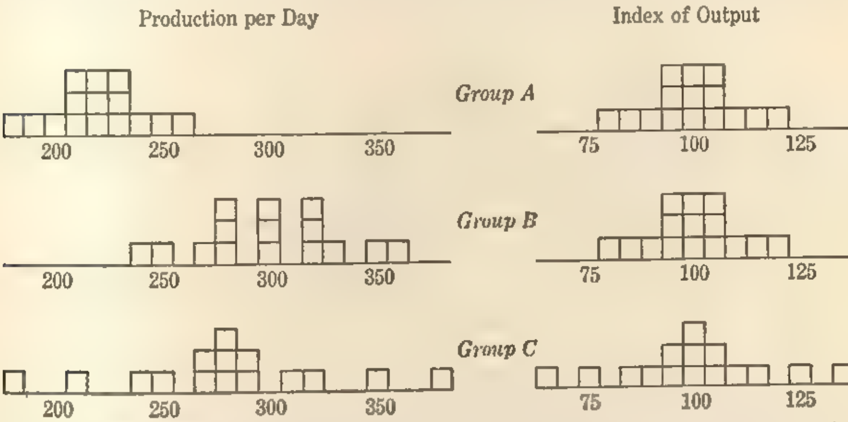


FIG. 3-1. Comparisons of distributions of rate of production and indices of output for three hypothetical groups of workers, *A*, *B*, and *C*, each group being assigned to a different model of the same machine.

used as bases for evaluating individual workers. Unfortunately, however, this is just the sort of situation where this type of index is most needed. In a department store, for example, sales vary from department to department due to variations in the selling situation and the nature of the goods offered for sale. Some procedure is needed to render their gross sales comparable. Since at most there will be only 15 or 20 persons in any one department, no great confidence can be put in the stability of the average of the sales in the different departments.

Second, the variability in the indices of output for the various groups must be the same. If the indices for one group of workers cluster closely around the average of 100, while in another group the indices deviate far below and above 100, then the output indices of these two groups cannot be considered comparable. In Fig. 3-1 the distribution of the indices of output of workers assigned to machines A and B have the same variability, while the distribution of indices of workers assigned to machine C is much more variable, containing more very high and very low indices. It might well be that the characteristics of machine C are of such a nature

that inferior workers are unduly penalized, and superior workers are correspondingly given an undue advantage. The best worker on either machine A or B can achieve an index of only 120, whereas one on machine C can achieve an index as high as 135.*.

Third, the workers in the various subgroups must be comparable in ability. If, over a period of time, the better workers were assigned to one group and the poorer workers to another group, then reducing their production rates to a comparable basis would penalize the superior workers and give an undue advantage to the inferior workers. The selective placement of workers in different groups on the basis of differences in their ability is common practice. In a factory the more able workers are assigned to the faster machines, and in a department store the better salespersons are assigned to departments where the goods are more valuable and the sales are likely to be higher. If any wide differences of this nature exist between two groups, the method of achieving comparable indices of output as described above should not be used.

Average output is not always a satisfactory standard of production. In certain cases it is unwise to use average production as a standard for the computation of efficiency indices, and in other cases such an average is of doubtful value. In jobs for which vigorous recruitment campaigns have been utilized in order to procure top quality workers, it would be unfair to evaluate their performance in terms of their own average production. Furthermore, if workers know that they are to be judged in terms of their own performance rather than by some independently determined standard, they might adopt the point of view that they would have nothing to lose and much to gain by reducing their output as a group. Average production as a standard is subject to variations in the motivation, interest, degree of training, and skill of workers, all of which contribute to the workers' performance in ways that lessen the value of such an average as a standard. Finally, in certain situations it is impossible to obtain an average. With new jobs, for example, any standard would have to be established on grounds other than workers' performance since records of job performance are nonexistent. Some jobs consist of a large variety of different tasks, each of which would require its own standard. In a repair shop, for instance, the operations are so numerous and varied that reliable average times to complete each task cannot be determined, and hence standards for each task would have to be set up on other grounds.

Estimates of Standards from the Performance of Specially Selected Workers. To meet the foregoing problems standards may be set on the

* Certain statistical procedures are available for rendering not only the averages of different groups comparable but also their variabilities. The most commonly used procedure is to compute standard scores.[12]

basis of the performance of specially selected individuals. These people might be foremen, pace setters, or other workers whose performance might be expected to be "uncontaminated." It is recognized, of course, that such individuals will not be representative of the ordinary workers on the job, and that in general their performance will be superior to the average worker. Hence these solutions involve a second step, the use of a "leveling" factor. The aim is to "correct" the data obtained from the specially selected individual so that a standard is set which describes an "average" worker.

A number of experimental studies have been made of the effectiveness of leveling procedures. The general approach is to present the same task, executed at several different known rates, to experienced time-and-motion-study men and have them estimate the speeds using an intermediate one as a standard. For example, in one typical investigation, time-and-motion-study men were trained to be able to judge accurately rates of speeds of 80 per cent, 100 per cent (standard), and 120 per cent for a variety of tasks in factory operations.[15] The experts were then tested for accuracy in judging other speeds. Two-thirds of the estimates were found to be correct within 10 per cent, and 89 per cent within 15 per cent. Thus, while there is some validity in an expert's judgment of speed, in a substantial proportion of instances the errors are large. One difficulty with investigations such as this is that they purport to indicate the accuracy with which a standard performance can be established, and yet the standard is chosen arbitrarily. The mere fact that an arbitrarily selected time can be estimated with a fair degree of accuracy by an expert does not mean that that particular time is the correct standard to use.

A further difficulty with these procedures arises from the fact that they have been conducted largely under artificial conditions. In one investigation in a factory situation it was found that the speed of operation of the most rapid worker was 67 per cent greater than that of the slowest.[7] When leveling factors were applied to compensate for estimated differences in skill and effort, the superiority of the best worker was still very large, being as much as 53 per cent greater than the rate of the slowest worker. If the correction factors had been adequate, the differences between the best and the poorest workers would have disappeared.

It can be concluded that attempts to equate for differences in skill and effort between the specially selected workers used in setting operation times and the general run of workers have not proved wholly successful.[11] While such corrections are somewhat better than no correction at all, standards which utilize them cannot be considered to have attained a high degree of accuracy.

Experience with Similar Tasks as a Basis for Setting Standards. If champions run the 100-meter dash in about 10.3 sec., and the 200-meter

dash in about 20.7 sec., then it might be expected that they would do a 150-meter dash in about 15.5 sec., or, allowing a little extra for the time consumed in the getaway, the time might be set at 15.6 sec. It is argued that in a similar fashion standards might be set for industrial tasks. If, in a garment factory, the standard for making a size 10 item is 16 min. and for making a size 14 item, 18 min., then it would seem reasonable to set a standard of 17 min., or perhaps 17.1 min., for a size 12 item.

This procedure frequently is used, either explicitly or implicitly, for the setting of standards. For some jobs it would seem reasonable to set standards for a new task in terms of present practices with other similar tasks. Obviously, the accuracy of the new standard set by this means will almost wholly be a function of the accuracy of the standards from which it is estimated.

Time Study. Time study is the traditional method of the industrial engineer for the setting of work standards. It has a long history which has been characterized by many heated disputes. Time study has been termed by some a nefarious device for speeding up work, and by others the ultimate in scientific methodology for setting work standards. As is usually the case with any controversial issue, neither extreme description is true. Any procedure can be misused, and none is as good as its most ardent proponents claim.

The purpose of time study is to set up standards for work, but specific statements of its objectives vary considerably. In general, the purpose of time study can be said to be the determination of the time an individual with ordinary qualifications, working with normal effort, should take to do a specified task. Any given worker can be evaluated in terms of the degree to which his work corresponds with this standard.

Time study is not a single method, as is generally believed. Rather, it is a general approach to the setting of times for work. Different industrial engineers support somewhat different methods. The processes essentially involve the following six steps: [1]

The fractionation of the task into its elemental parts, each of which is timed

The determination of those elements which are essential for completion of the task

The determination of an operation time for each element by selection or correction of the original data

The determination of the operation time of the total task by adding together the operation times of all the elements

The determination of extra time allowances

The determination of the standard time for the task by adding together operation time and extra time allowances

The task for which the standard is to be set is subjected to a time-and-motion analysis such as was described in the previous chapter. This involves a fractionation of the task into its basic parts or elements, and the

timing of each. The choice of the actions to be timed always poses a problem, and various types of solutions have been adopted. The solutions involve a consideration of what the time-study man considers to be the essential motions required to complete the task, a selection from the original data collected of a "true" or "representative" time for each motion, and an adjustment of these times by a leveling procedure for differences in the skill and effort of the workers being studied.

It is argued that any given worker is unlikely to perform the job in the way it ought to be. Hence from an examination of the pattern of motions revealed by the motion analysis the time-study man sets up an ideal or best set of motions. If it appears to him that the sequence of motions actually made by the worker studied is adequate, he may adopt them as they stand.

A given worker will vary somewhat from time to time in the rapidity with which he performs a given element, and there will be some variation among workers. The average time may be taken as the standard, but ordinarily the times actually taken by some workers to complete an element may be regarded with suspicion. Hence some times, such as those that are very long or very short, may be discarded as being abnormal. The shape of the distribution of times leads to judgments concerning the worker and leads to certain of the times being accepted and others rejected.[18] Suppose that in a relatively few instances a worker performs quickly but most of the time he is relatively slow. The inference is drawn that the worker has poor motivation. The few instances of short times would be retained as being representative of what the worker can do when he utilizes proper effort, and the rest of the data are discarded. If the situation is just the reverse, that is, there are many short times and few long times, it would be concluded that the worker was too highly motivated during the time-and-motion study. The short times then might be given less weight. If the worker shows considerable variability in the speed with which he performs an element it would be concluded that he has not developed skilled habits of work, and hence only the shorter times would be used in determining the standard.

Similarly, if three workers were timed, and the times of two were very similar while those of the third were quite different, data obtained from the latter worker might be discarded. In making judgments such as these, the time-and-motion-study man is likely to obtain all the information he can concerning the skill, motivation, etc., of the workers studied, and use it as a basis for the selection of a particular time as the standard for each element.

The operation time for the total task is found by summing up the times set for each of the elements. Leveling factors may be applied if the workers used in the study are not considered representative. Finally, extra

time allowances are added to give the time standard set for the total task.

It should be apparent from the foregoing description of time study that it is far from an exact method.[11] The subjective judgment of the time-study man enters in at many points, and he has few established rules to follow in cases of doubt. The choice of operation times of the elements is clearly arbitrary and the factors that influence this choice vary from one situation to another. Inferences concerning the motivation, effort, and skill of the worker based upon the shapes of the distributions of element times result from unproved assumptions about ways in which people behave. The arbitrary character of extra time allowances has been discussed earlier. A further limitation of time study is its acceptance and use of the untenable conception that human behavior is merely the simple sum of elemental movements. This notion will be examined in more detail in a later chapter.

The final conclusion that can be drawn concerning time study is that it simply does not possess the precision and objectivity claimed for it. In certain respects it is superior to other ways for setting standards for work, since it does attempt to correct for biasing factors. In effecting these corrections, however, there is considerable possibility for the introduction of other types of error. The fact that final standards, even when set by trained and experienced time-study men, are subject to review and arbitration further supports the view that such standards are not to be interpreted as possessing "final" or "absolute" meaning in any sense.[4, 11]

Experimentally Determined Times as Standards. This procedure is based on the fact that a given specific motion may be required in the performance of many different tasks. Once the time required for a given specific motion is known it can be applied to new tasks in which the motion is present and for which a standard is needed. Experimental times would have to be worked out for a large variety of specific motions. Establishing a standard for a new task would require finding the particular specific motions involved in the task and summing the times established for these reactions.

Suppose in an assembly operation the worker used such elementary motions as reaching forward with the right arm to grasp a small object, moving the right arm and hand to the center of the workspace, moving the left arm slightly to the left and forward to grasp an object, moving the left arm and hand to the center of the workspace, etc. The time for each of these motions—and for others—would have been worked out separately. The standard for the assembly task could then be obtained by selecting the appropriate elemental motions and summing the times required for their execution.

This procedure is subject to most of the limitations already mentioned for time study. Certainly the times found for the specific motions would be a function of the training, experience, motivation, and other personal characteristics of the workers studied. Many subjective judgments of the job analyst are required in the process of selecting the specific motions and in determining the appropriate times for them. The method rests on the fallacious assumption that a complex task can be obtained from a simple summing of elemental movements. Except for convenience, the method has no advantage over the method of time study.

Rational Analysis as a Basis for Setting Standards. In some situations standards are set by experts or interested parties. Usually such people arrive at a conclusion as a result of discussion, bringing together such pertinent information as the results of time studies, production records, and reports by workers and supervisors. The attempt is to integrate this information and arrive at some standard. This process has been termed rational analysis, presumably because it involves the exercise of judgment on the part of sincere persons. It is probably more accurately described as intuitive, since all of the factors that influence the final decision are neither recognized nor explicit. To be sure, the final decision can be justified after it has been rendered, but this is more rationalization than rational analysis.

Attempts to set standards of work through rational analysis point up the heart of the whole problem. Actually what is being sought is an answer to the question of how much work a normal individual ought to do—that is, what is a fair day's work? The question has no relation to how much work people do do or how much they can do, but to some ideal. The type of judgment called for is not one of fact but of value.[3] Values derive from attitudes, needs, and desires, which in turn arise out of the particular culture at the given time. The origins of values can be studied, and the values themselves described, but their goodness can be judged only in terms of the degree to which they satisfy the interests of a particular group. A work of art is not inherently good or bad, rather its qualities are evaluated with respect to how well they correspond to the values of art critics, the general public, or some other particular group. Similarly, a standard of work which requires the individual to apply himself with great effort for 10 hr. a day and sends him home exhausted is bad only if the values of the society say it is so. Again, if the standards permit the individual to work at a leisurely rate and with little effort, such standards are good only if the results achieved are commensurate with declared values.

OTHER MEASURES OF PROFICIENCY

Work-sample Tests. When working conditions are such that measures of output are difficult or impossible to obtain in the plant or office, work-sample tests may be used to get a measure of productivity. Sometimes output records of employees are not useful measures of performance because they are not accurately kept. Sometimes the worker's performance is difficult to register accurately, as in the case of the production of a file clerk in a busy office. Sometimes the process of observing and registering employee performance interferes with the rate of production. For example, recording all of the items of work of a machine operator may slow him down and an inaccurate rate of work would be described. Work-sample tests overcome many of these handicaps and make possible an accurate measurement of the worker's production under a standard set of conditions.

In a work-sample test the employee performs exactly the same work as that involved on the job, either for a fixed and limited time, or until a given number of items of work are completed. Perhaps the most commonly employed work-sample tests are the typing proficiency tests wherein the typist transcribes standard material, and speed and accuracy of performance are recorded.

Obviously, the usefulness of work-sample tests will be in direct proportion to how well they represent the actual work on the job. If important aspects of the job are not reducible to the test situation then the test will not be representative. If the test situation is overly simplified in order to get more objective measures, or in order to facilitate the collection of performance records, it may cease to measure the functions desired for the production standard. For such jobs as automotive repair, it might be impossible to present standard work tasks to be accomplished, and hence the work-sample test would be of questionable value.

One of the important prerequisites of a work-sample test is that all of the individuals taking it should be subjected to exactly the same situation and task. It would be unsound to test one bookkeeping machine operator in a busy office with double-entry problems for 2 hr. and another in a quiet room on single-entry problems for 15 min. One of the obvious advantages of the work-sample test is that it can be conducted under uniform conditions. Extraneous and unimportant factors can be brought under control and accurate records of performance obtained. This often requires the test to be conducted in a separate room, in some segregated part of the plant, or at a machine or workplace especially arranged, so that output can be measured without interfering with the

testing procedure, or the testing activity interfering with normal production schedules.

Length of Service. In many situations the capacity of the individual to continue on the job is an important measure of his success. Certainly if the organization must spend considerable time and money in training, an individual who completes the training and shortly thereafter leaves the job must be considered an unsuccessful employee.

Length of service probably reflects the individual's ability to adjust himself to the working conditions, his ability to get along with his fellow workers, his satisfaction with the job, and similar factors. In addition, it may measure to some degree his ability to perform the work involved in the job, since inferior workers will be detected and released or transferred to other jobs. However, the relative importance of each of these factors in determining the length of time the employee stays on the job certainly will vary markedly from job to job. An exit interview directed to the discovery of causes of separation is essential if all the facts are to be ascertained.

Another value in using length of service as a measure of job success arises from the fact that as an individual continues on his job he increasingly develops skill and knowledge. Studies of learning indicate that where proficiency is measured during long periods of practice improvement continues to be manifest. "Old hands" have experienced situations that arise only occasionally, and are better prepared to cope with them than is the new worker. The worker who stays on his job for a long period of time is prepared to contribute to his organization in many important ways.

Amount of Training Necessary. The nature of many jobs is such that the organization must train the worker for them. Training time can be used as an index of worker proficiency. The individual who takes a long time to finish such training, by and large, cannot be considered as good as one who learns in a relatively short time. Success during the learning period can be used as an index where the training period is expensive to the organization, where the trained employees are relatively homogeneous in respect to job proficiency, or where training time is related to later success.

If training time is used as an index of the worker's potentiality for later success on the job, it is important that the relation between these two factors be carefully determined. Since the employees being trained usually do not have the same past experience related to the job for which they are in training, any such related experience will tend to produce individual differences in the required training time. If this experience can be measured and taken into account in evaluating the

length of time taken to complete training, a more accurate appraisal of training time as a measure of subsequent success on the job will be attained.

Ratings by Supervisors. The opinions of foremen, managers, and others in supervisory authority are used as measures of job success. Such opinions are recorded on some kind of rating form. Rating of workers by their supervisors has enjoyed wide use, and frequently has been abused. The various types of rating techniques and the problems involved in making ratings will be discussed in the next chapter.

FACTORS INFLUENCING MEASURES OF PROFICIENCY

A measure of job success can be valid and reliable and still not be wholly adequate. Extraneous factors influencing the worker's performance, but not reflected in the validity and reliability of the measuring device, may introduce errors. Some of the more important factors of this kind are considered below.

Experience of the Workers on the Job. On most jobs there will be a correlation between the length of time on the job and job proficiency,

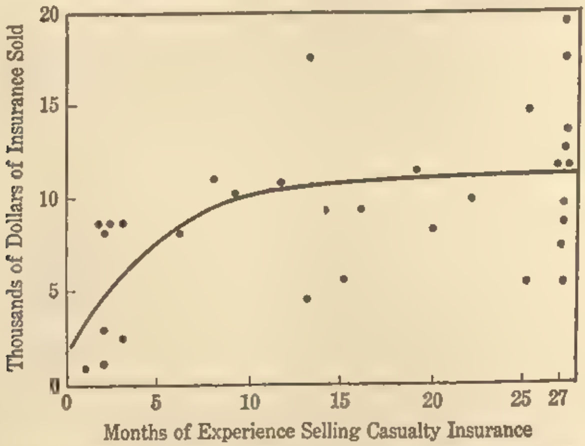


FIG. 3-2. The relationship between monthly sales and amount of experience of 31 casualty insurance salesmen.

the newer workers being less productive than the older ones. In Fig. 3-2 is shown the relationship between the monthly production of 31 casualty insurance salesmen and the length of their experience selling for their particular company. There appears to be a rapid increase in productivity up to about 10 months with a very slight increase thereafter. It is clear that the productivity of these salesmen can be compared only when their experience is the same or when it is equalized in some manner.

Since it is often desirable to compare the effectiveness of work of different employees on a job, it is advantageous to correct for the experience variable. One means is to use a trend line indicating the relationship between experience and production. The relationship between amount of experience and the measure of proficiency is represented graphically, and the trend line best representing the relationship is drawn. Each individual's proficiency is then computed in reference to this line. In Fig. 3-2 it will be seen that the average production (as given by the trend line) is $5,500 for salesmen with 3 months' experience, whereas for salesmen with 20 months' experience it is $11,000. A salesman with 3 months' experience whose production is $8,500 is $3,000 above the average for his experience level, while one with 20 months' experience whose production is $8,500 is $2,500 below the average for his experience level. With experience taken into account, the first individual is obviously superior to the second as a salesman of casualty insurance.

Age. Age is another factor that might distort the measurement of proficiency. In some cases the performance of older workers will differ from that of younger ones. The effect of age may be corrected for in the same manner as experience. In many instances, however, age and experience will be highly related, so that the correction for the one will automatically correct for some of the effect of the other.

Sex. It is not unusual to find that on a given job members of one sex will be superior to members of the other. Under these circumstances some adjustment may be necessary. A very simple solution to the problem is to compare the performance of a given individual with the performance of like-sex individuals. The individual's relative superiority is expressed in terms of how far he or she deviates from the average job performance of individuals of the same sex.

The Time Interval during Which the Measurements Are Made. In many instances job performance varies with time of day, day of week, and season of year. Thus the time interval during which job proficiency is measured will be a further variable that may distort comparisons between workers unless it can be shown that each worker's relative standing on the job is not affected by the time of day, day of week, and season of the year.

It may well be expected that, with differences in the time variable, one person's performance will vary systematically from that of another's. For example, the production of a new sales clerk, hired during the Christmas season, will be higher than that of an experienced sales clerk during the summer period. The superiority of one over the other is not the result of differences in ability, but of differences in ease of selling because of seasonal changes in the market. Correction can be made by

comparing each individual's performance with the average of workers doing comparable work for the period of time under consideration.

COMBINING SEVERAL MEASURES OF PROFICIENCY INTO A SINGLE COMPOSITE MEASURE

The discussion thus far in this chapter may leave the reader with the erroneous impression that there is always one measure of job success, or at least one best measure. This is seldom true. For example, determining the relative success or failure of streetcar motormen may require that all of the following variables be taken into account.

Number of collisions with vehicles
Number of collisions with pedestrians
Number of traffic violations
Number of commendations received from the public
Number of complaints received from the public
Number of times company rules are broken
Number of sleep-overs (tardiness in reporting)
Number of times schedules are broken
Number of reprimands received from inspectors
Ratings by inspectors
Errors reported by dispatchers

It is apparent from the foregoing list that the success of motormen can be measured in a variety of ways, no single way giving an adequate index of success on the job. In order to classify workers on the basis of a single measure of general job success, these separate measures must be combined into a single one. This means that a weight must be assigned to each of the component measures of proficiency.

Equal Weighting. Allowing the various components to carry equal weight in a composite criterion is not a satisfactory procedure, since seldom will the components be of equal importance in job success. Arbitrary assignment of equal weight to the various measures merely ignores the problem rather than solves it. In the following paragraphs consideration will be given to three methods that have been devised for determining weights to apply to different criterion measures in order to combine them into a single composite criterion. In addition, two other procedures for effecting a combination of several measures of proficiency will be discussed.

Weighting According to Judgments by Experts. The simplest procedure is to have each of the component measures of job proficiency judged or rated according to its importance by experts, *i.e.*, persons who are very familiar with the job in question. In a discussion group foremen, supervisors, or other individuals who are acquainted with the job consider all of the measures, and express opinions concerning their relative im-

portance. For practical purposes the judgments by experts would seem to be the best method of arriving at weights for the component measure. Important information, such as the causes of separation of employees, and the parts of the work in which employees are particularly effective or deficient can be utilized. The composite index would have considerable meaning in terms of the actual job performance.

The primary difficulty with the use of experts' opinions is that satisfactory agreement in their ratings is seldom attainable. To illustrate how such differences among experts lower the dependability of the estimated relative importance of criteria, the results presented by Stead and Shartle are given in Table 3-4.[16] Each of six experts rated the importance of a

Table 3-4. The Correlation of Each of Six Experts' Ratings of the Importance of a Number of Criterion Measures with the Ratings of the Remaining Five Experts

Expert	*Coefficient of correlation*	*Expert*	*Coefficient of correlation*
A	.20	D	.14
B	.26	E	.20
C	.34	F	.22

number of proficiency measures on a scale from 0 to 9, and the correlation between the ratings of each expert with the average of the other five was calculated. From the data shown in Table 3-4 it will be observed that the correspondence between experts' judgments is relatively low, which is not at all uncommon for this type of rating. In another situation, for example, experts rated the importance of ten measures of proficiency for each of three technical jobs. The correlations between the ratings of importance made by experienced job analysts for the three jobs were .60, .41, and .34, respectively.

Weighting According to the Reliability of the Criterion Measures. Another method, which has been used infrequently, is to weight the component measures of proficiency in terms of their reliability. The argument in favor of this procedure is that less reliable measures should not be given as much weight as more reliable ones, since they contain a larger proportion of errors of measurement. However, the validity or representativeness of the criteria should also be considered in the determination of their relative weights in the composite. A given measure may be highly reliable and consequently would be given a high weight, but still be relatively unimportant because it is not highly representative of the job requirements. If the measures available are about equally representative of the job, then weighting them in terms of their reliability is a very effective procedure.

Weighting According to an Assumed Underlying Variable of Job Success. Another method for combining different measures of proficiency into a composite is based on the assumption that all of the component measures are really indices of some single ability, or various aspects of a single ability, which underlies performance on the job.* In essence it is assumed that underlying all measures of success on a job there is a common factor. Any given index of job success is thought of as including two kinds of factors, those which truly measure job success and those which do not. The former are considered to be important since, pursuant to the basic assumption in the method, they are involved in all indices of success for the particular job. The latter are not considered important since they have no generality. On the basis of this assumption it is argued that each measure should be weighted in proportion to how highly it is correlated with this underlying common factor.

It follows from this assumption that the intercorrelations between different measures of job proficiency will necessarily be greater than zero, and for the assumption to be meaningful the intercorrelations should be at least moderately high. That these conditions are not always fulfilled may be seen from data presented by Stead and Shartle, which are reproduced in Table 3-5.[16] The average of these correlations is relatively low, being .38. It is significant that the correlation between number of sales and returns, possibly the two most important of the measures in the group, is only .01.

Table 3-5. Intercorrelations between Various Indices of Job Success of Department-store Salespersons

Criterion	Returns	No. of Sales	Quota
Gross sales	.58	.47	.65
Returns		.01	.32
No. of sales			.24

Multiple Cutoffs. Another procedure for combining different measures involves a series of multiple cutoff points. For each measure of proficiency a minimum standard is set, and a worker who is above the standards on every measure of proficiency is deemed successful, whereas a worker who falls below on any one of the measures is considered un-

* The mathematical procedures involved in this method have been discussed by Edgerton and Kolbe.[9] It is interesting to note that this weighting procedure gives maximum differences between individuals in terms of composite criterion scores and minimum differences within an individual in terms of his component criterion scores.

successful. On this basis workers are classified into two groups, the successful and the unsuccessful. With this procedure the point of view is adopted that, even if a worker is superior in all but one aspect of the job, every single aspect is considered essential, and failure in any one is tantamount to complete failure.

With certain jobs this procedure makes a great deal of sense. As an example, no matter how good an airplane pilot is in take-offs, navigation, cross-country flying, etc., if he is poor in making landings he must be considered unsuccessful. The multiple cutoff procedure clearly is akin to the critical incident technique of Flanagan described in the previous chapter.[10]

The "Dollar Criterion." Brogden and Taylor have suggested that the concept of cost accounting can be applied to the problem of combining different measures of job proficiency.[6] The argument is that the contributions that a worker makes to an organization and the costs that he incurs for it can be expressed in terms of dollars and cents. The contributions he makes will be in terms of the dollar value of his output. Similarly, the costs that result from the worker's activities through wastage, accidents, errors, etc., can be expressed in terms of dollars. Finally the net contribution in terms of dollars made or lost for the company can be determined. Since all measures of performance are expressed in the common term of dollars they are readily combined.

This approach to the problem of evaluating the effectiveness of workers will appear to many to be the most pertinent and meaningful. It does, however, have some important deficiencies. Cost-accounting procedures are expensive, especially when the values of many different factors need to be individually reckoned. Furthermore, such procedures are by no means exact and it is often difficult to assign the cause of a gain or a loss and hence it must be done arbitrarily.[13] Finally, it is obvious that in many instances the contributions and costs that a worker may incur for an organization may not be measurable. For example, how can the dollar value of the efforts of an elevator operator be determined? Or the good will gained or lost to a store by a sales clerk? On the other hand, this "dollar criterion" is valuable in the sense that it will tend to focus attention on the development of objective indices of job proficiency. It also emphasizes the fact that a worker's contribution to an organization has many aspects, all of which should be taken into account if an accurate appraisal of his contribution is to be made.

REFERENCES

1. Barnes, R. M.: "Motion and Time Study," Wiley, 1949.
2. Bellows, R. M.: Procedures for evaluating vocational criteria, *J. Appl. Psychol.*, **25**, 499–513, 1941.

3. Benne, K. D., and G. E. Swanson: Values and the social scientist, *J. Soc. Issues,* **6,** No. 4, 1950.
4. Blum, M. L.: "Industrial Psychology and Its Social Foundations," Harper, 1949.
5. Brogden, H. E., and E. K. Taylor: The theory and classification of criterion bias, *Educ. and Psychol. Measmt.*, **10,** 159–186, 1950.
6. Brogden, H. E., and E. K. Taylor: The dollar criterion—applying the cost accounting concept to criterion construction, *Personnel Psychol.*, **3,** 133–154, 1950.
7. Cohen, L., and L. Strauss: Time study and the fundamental nature of skill, *J. Consult. Psychol.,* **10,** 146–153, 1946.
8. Cook, H. E., and G. E. Manson: Abilities necessary in effective retail selling and a method of evaluating them, *J. Personnel Research*, **5,** 74–82, 1926.
9. Edgerton, H. A., and L. E. Kolbe: The method of minimum variation for the combination of criteria, *Psychometrica*, **1,** 183–187, 1936.
10. Flanagan, J. C.: Critical requirements: a new approach to employee evaluation, *Personnel Psychol.,* **2,** 419–425, 1949.
11. Gomberg, W.: "A Trade Union Analysis of Time Study," Science Research, 1948.
12. Guilford, J. P.: "Psychometric Methods," McGraw-Hill, 1936.
13. Knauft, E. B.: A selection battery for bake shop managers, *J. Appl. Psychol.*, **33,** 304–315, 1949.
14. National Research Council: "Fatigue of Workers," Reinhold, 1941.
15. Schell, H. A.: A study of effort rating, *Modern Mgmt.*, **9,** No. 1, 19–20, 1949.
16. Stead, W. H., and C. L. Shartle: "Occupational Counseling Techniques," American Book, 1940.
17. Toops, H. A.: The criterion, *Educ. and Psychol. Measmt.*, **4,** 271–297, 1944.
18. Wiberg, M.: "Work-time Distribution," McClure, Hadden, and Ortman, Chicago, 1951.

CHAPTER 4

Rating Methods

Appraisals in the form of personal estimates are needed in solving many personnel problems arising in industry. Evaluations of this type, termed ratings, are used to gauge the proficiency of workers in the performance of their jobs, their progress in training, and their potentiality for promotion. These ratings, which are quantitative statements of opinion, can be made in a variety of ways. Some of the procedures communicate adequately the ideas of the raters while others do not. In this chapter consideration will be given to the various ways in which personal appraisals can be expressed and recorded. Since ratings are nothing more or less than expressions of opinion, they are subject to all the errors characteristic of human judgment. It is necessary, therefore, to consider the most common kinds of errors and to evaluate the procedures that can be used for overcoming or minimizing them.

VALIDITY OF RATINGS

As with other psychological measuring instruments, the problem of validity is a basic one in the development of rating procedures. It should not be slighted just because its solution appears more difficult than with most other measures of proficiency. It is important that ratings be valid in the sense of being accurate assessments of the worker behavior being evaluated.

The Subjective Nature of Ratings. The common notion that ratings can be objective descriptions of fact must be regarded with extreme caution. Those who believe that traits such as leadership, persuasiveness, and politeness can be rated objectively need to be disabused. There is no such device as an objective rating procedure. It is possible, of course, to record certain events that occur, such as the behavior an individual displays under a given set of circumstances. Counting the number of times an individual behaves in a given way when a given set of circumstances arises may be a perfectly acceptable way of evaluating workers but it is not rating. Rating involves personal estimates and subjective

judgments. The number of times a subordinate addresses his superior as "sir" is one way of evaluating politeness. The superior's personal opinion of the subordinate's behavior as expressed in a rating is another. They may be equally good indices of politeness and still be unrelated. Their validity as indices will depend upon the nature of the behavior which each is measuring.

One of the reasons for the widespread use of ratings in industry has been the failure to develop objective techniques for measuring complex types of behavior. Through subjective estimates, ratings make possible a quasi-quantification of these complex behaviors. The validity of these ratings needs to be determined by comparing them with more objective measures. But here there is a dilemma. Ratings are used because such objective measures are not available or are very difficult to obtain. It is then doubly necessary that the factors known to affect the validity of ratings be thoroughly investigated. Only then can conditions be set up which will be conducive to producing the highest possible validity in subjective types of measures.

Validity as a Function of the Nature of the Behavior Being Rated. Ratings are procedures for quantifying the opinions of one person about the behavior of another. One of the factors affecting the validity of ratings is the complex nature of the behavior being rated. On the one hand, it may be desirable to have a rating of the amount of production of a worker at a bench lathe. The judgment can be based upon direct observation of the items produced by the worker at the completion of stated work intervals. On the other hand, a measurement of the leadership ability of a worker may be desired. Here is an intangible kind of behavior, complex in nature, having few manifestations which can be clearly recognized and noted on successive occasions. It is to be expected that more agreement is possible between a judge's estimate of a person's behavior and his actual behavior in the simpler situation than in the more complex, intangible type of behavior situation. Validity of ratings then will be a function of the degree of complexity of the behavior being assessed.

In the two extreme types of situations mentioned above the need and importance of measuring instruments are just the reverse of the degree of validity or agreement thus far obtained. That is, the greatest need in psychological measurement is for measures of complex behavior such as attitude, loyalty, or leadership. In this area the validity of the instruments has been low. The problems of measurement are less difficult in the areas where the behaviors are more simple and can be assessed by such measuring devices as production indices, test scores, etc. Here several substantially valid instruments have been perfected.

Sometimes complex traits are considered composites of a number of

simpler traits. It has been pointed out that, in so far as possible, ratings should be made of the simpler traits rather than of the more general traits.[6, 16] When workers are rated on a general trait there is greater possibility of error since different raters will base their judgments on different aspects of performance included under the trait name. Furfey found that agreement among raters is increased if the trait being rated is carefully analyzed.[7]

It is generally believed that those traits which are of the so-called objective type are more accurately rated than those of a subjective nature. It is not an easy matter, however, to differentiate traits in terms of objectivity. What is it that makes one trait objective and another subjective? Examination of the evidence indicates that, in many cases, traits that are generally considered to be more objective than others actually are less consistently judged. Thus punctuality would be considered by most persons to be more objective than cooperativeness; yet Slawson found the reliability of punctuality to be .41, while that of cooperativeness was .52.[22] On the basis of results thus far obtained the best conclusion that can be drawn appears to be that of Symonds, viz., that the traits that are most consistently judged are those which influence external events or leave their mark on things.[25] In other words, traits for which there is external and objective evidence are the ones that are most accurately judged.

Validity as a Function of the Rater's Knowledge of the Behavior to Be Evaluated. When complex behavior is to be evaluated valid ratings cannot be expected if the rater does not understand what it is he is to assess. The subjective notions of the rater make a significant contribution because they will affect the interpretation and meaning which he will give to the behavior. The less well defined the behavior the more opportunity there will be for error since different raters will interpret the meaning of the behavior differently. The behavior to be assessed should be carefully defined with all of its aspects specifically described.[13]

Usually when complex behavior is to be rated, the judge is dependent upon calling up in memory instances of the worker's behavior that he has observed in the past. The judgment of the rater about the individual being rated is in terms of how the behavior of the latter is perceived, interpreted, and remembered by the former. The reaction of the rater is a purely personal experience and is likely to be different in different raters. If a rater has a distorted notion of the trait or characteristic being evaluated, the situations he will call up will not contain the exact behavior wanted. The ratings then cannot have high validity.

It is essential that special instructions be given the raters concerning the nature of the characteristic to be rated. This requires more than a general description in which the more specific meanings are omitted and

their insertion left to the rater. It will be in terms of the specific meanings that accurate ratings will be attained. These specific meanings should be supplied to the raters, and examples of concrete expressions of the behavior should be described so the raters will get an accurate picture of what is to be evaluated. Such definition and specification will tend to focus the attention of all raters on the same areas of behavior for all of the individuals to be rated.

Validity as a Function of the Rater's Knowledge of the Behavior of the Worker Being Rated. It is obvious that the more familiar the rater is with the behavior and characteristics of the person to be rated, whether through greater opportunity for observation or through longer time of observation, the more accurate will be the ratings.[4] Just what constitutes adequate familiarity will have to be determined for every job. For certain simple routine jobs, one or two months will suffice, but for more complex and varied jobs a period of years may be necessary.

Even though an individual acts as the supervisor of a group of workers, it does not necessarily follow that he is sufficiently familiar with them to make accurate ratings. In many instances the supervisor has operational functions of his own and only incidentally performs personnel duties and evaluative functions. Obviously the length of time an individual supervises a group of workers may not accurately indicate the extent to which he is familiar with their performance. Long-time casual observation is probably less effective than short-time intensive observation.

Validity as a Function of the Bias of the Rater. Even with knowledge of the behavior to be measured and knowledge of the worker's behavior to be rated, valid ratings will not be obtained if the judgments of the rater are influenced by personal bias. Here the concern is with the introduction of inaccuracy into the ratings because of personal notions or experiences of the rater which prevent him from making accurate judgments. For example, it has been reported that raters assign higher ratings to their colleagues and fellows than they do to other acquaintances.[17] This means that an individual cannot be expected to rate accurately any close friends who are working under his supervision. In those situations where the supervisors and the workers form a tightly knit social group, valid ratings by the supervisor cannot be expected.

Validity as a Function of the Purpose of the Ratings. Ratings which are intended to be measures of worker proficiency should not also be expected to serve the purpose of motivating the worker. If the rater has in mind this dual function then the validity of the rating as a proficiency index will suffer.

Supervisors often use ratings as motivational devices. If it is felt that a particular worker needs bucking up, depending upon the personality

of the worker, a supervisor might rate him either lower or higher than his true performance merits. In effect the supervisor is using reproof or praise in an attempt to motivate the worker to higher levels of performance. If ratings are used for such a purpose they are useless for evaluation. In other cases a supervisor might rate an individual higher than he ordinarily would if he felt that it would help the individual obtain a promotion or raise. This again makes the rating an inadequate estimate of the worker's ability or effectiveness.

Validity as a Function of the Amount of Time Available for Making the Ratings. In order to accomplish accurate ratings the supervisor must have time to collect, systematize, and analyze all of the impressions, opinions, and facts he has about each of his workers. The less time he has to organize his thoughts and materials, the less accurate will be his ratings. This factor is pertinent for validity since supervisors also have many operational duties to accomplish, and ratings ordinarily are an additional task fitted in as time permits—and usually there is very little time. Furthermore, many organizations call for reports of ratings on very short notice. Under such conditions the supervisor can only make perfunctory ratings, and they will be highly inaccurate. A supervisor should be allowed a minimum of half an hour of free time for the rating of each individual.

Validity as a Function of the Characteristics of the Raters. Guilford has made a careful review of the literature concerning the characteristics of individuals who are good raters.[11] In so far as the industrial situation is concerned, a good rater is one who:

Is interested in the ratings that he makes
Takes his time in making his ratings
Is well adjusted personally
Is sympathetic and understanding of people
Is not necessarily self-consistent
Is well aware of his own capabilities and limitations
Is trained and experienced, so that he understands the operations of, and compensates for, the errors that may affect accuracy of rating

As Guilford has pointed out, it cannot be said that there is a general judicial capacity. An individual may be a good rater for certain traits or for certain groups of workers and not for other traits or for other groups of workers. Unfortunately the capacity of individuals to rate others is not accurately appraised by the opinions of their superiors.[5] Raters judged as being good raters by their superiors differ very little in accuracy from those judged as being poor raters.

There are no thoroughly systematic studies dealing with the training of raters. Rationally it would seem that raters should understand the basic facts of individual differences among people, be informed of the

characteristic errors found in ratings, and have substantial information concerning the nature and requirements of the jobs held by the persons they are rating. It would also seem desirable to have them make at least their first ratings under the supervision of persons experienced in the problems of rating.

RELIABILITY OF RATINGS

The reliability of any measuring device is determined by observing the degree of consistency between independent and comparable applications of the instrument.

The Independence of Ratings. In order to achieve independence of ratings it is necessary that the ratings be made by different individuals. Successive ratings made by the same individual cannot be considered independent, inasmuch as personal biases, or any other constant error influencing the first set of ratings, would be likely to influence the second set of ratings. In addition the factor of memory prevents the successive ratings made by the same individual from being truly independent. It is apparent, therefore, that the reliability of ratings must be determined by observing the degree of consistency of ratings made by different individuals. This does not mean that the consistency of ratings made by a single individual is an unimportant consideration. Obviously, if a person is not self-consistent in his ratings of the same persons or things, his ratings will not be consistently related to those made by another individual.

The Comparability of Ratings. The factor of comparability poses a difficult problem. In the evaluation of workers it frequently is difficult to secure ratings by two or more individuals who have had equally good opportunities to observe the workers. Ordinarily the quality of the work of a group of employees is best known by their immediate supervisor, while the knowledge of their effectiveness gained by supervisors at higher levels generally is fragmentary and superficial. Only if all supervisory personnel concerned are equally familiar with the workers and their activities can their ratings be considered comparable.

The Basis of Unreliability of Ratings. Considerable evidence has accumulated relative to the causes of disagreement among raters.[11] The indication is that raters disagree primarily because they observe the individuals to be rated in different situations and under different conditions, and because they use different criteria for judging the same trait or characteristic. It follows from this evidence that reliability of ratings can be considerably increased by having the raters observe the individuals under similar situations, and by providing techniques for making the

ratings that will increase the likelihood that the traits or characteristics being judged will be evaluated on the same bases by all raters.

TYPES OF ERRORS MADE IN RATINGS

Ratings and rating procedures have been the subject of considerable study. As a consequence, the major types of errors made by raters are well known. The discussion here will be concerned with the two most common types, namely, constant errors and the halo error.

Constant Errors. Inspection of sets of ratings made by different persons will reveal certain individual tendencies on the part of some raters to

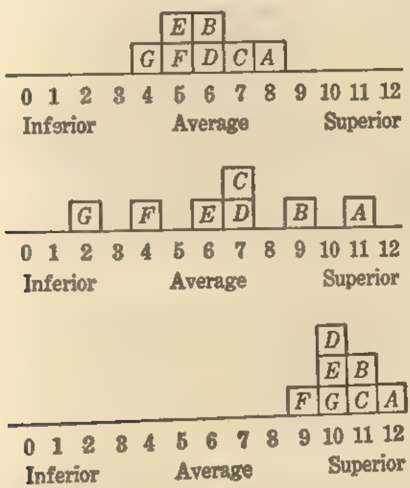


FIG. 4-1. Examples of constant errors in ratings of seven individuals as made by each of three raters.

differ from other raters in a constant manner. The two errors of this kind that are most frequently encountered are the *leniency error* and the *error of central tendency*. The leniency error refers to those situations in which the rater tends to assign high ratings to most individuals. Raters who tend to avoid rating individuals either very high or very low, and whose ratings tend to cluster closely around the average or central point on the scale, are said to be subject to the error of central tendency.

Examples of these errors are given in Fig. 4-1, showing the ratings of seven persons made by each of three raters. Although there is fairly good agreement among the three raters as to which individuals are relatively the best and which the poorest, the absolute ratings vary markedly. Thus individual A, who all raters agree is the best of the seven individuals, is rated 11 by rater I, 8 by rater II, and 12 by rater III. Rater I

shows no marked constant error since his ratings are distributed almost throughout the entire scale. By contrast, the ratings of each of the other two judges are closely grouped. Rater II has clustered his ratings closely around the average point on the scale, whereas rater III has clustered his at the high end of the scale. Rater II shows the error of central tendency, and rater III the leniency error.

The Halo Error. Another error which many raters manifest is called the halo error. By halo effect is meant the influence that a rater's general impression of an individual has upon his ratings of that individual on specific traits. In other words, ratings of some particular characteristic of an individual are contaminated by the rater's notion of the general worthiness of the individual. Thus a worker who is judged by a supervisor to be generally unsatisfactory will be rated low on all aspects of his job performance, even though he may be satisfactory in at least a few aspects. Similarly, another worker who, on the whole, is believed to be slightly better than average, will be given ratings slightly above the average on all phases of his job performance, even though he actually may be deficient in certain phases.

The ratings of two insurance salesmen made by two sales managers are shown in Table 4-1. Both salesmen were rated on an 11-point scale

Table 4-1. Ratings of Two Insurance Salesmen Made by Two Sales Managers, Showing the Halo Effect in One Set of Ratings (Sales Manager I)

Qualifications	Sales manager I		Sales manager II	
	Salesman A	Salesman B	Salesman A	Salesman B
Ability to sell new policies	8	3	9	1
Ability to obtain renewals of policies	8	3	6	4
Quality of services to clients	7	3	5	3
Quality of services to firm	8	3	5	2
Over-all value to firm	8	3	7	2

for each of five characteristics. The sales managers agree on the superiority of salesman A, but sales manager II has made a more discriminating analysis of the salesmen indicating in his ratings their relatively strong and weak points. The ratings of sales manager I are uniform for each salesman and clearly show the halo effect.

Since a rater who is subject to the halo error tends to give uniformly the same rating to each individual on all characteristics rated, the intercorrelations among the ratings are quite high. In Table 4-2 are given

Table 4-2. Coefficients of Correlation between Ratings of 26 Girls on Five Aspects of the Job of Inspector-Packer of Biologicals as Made by a Forelady

Job aspect	2	3	4	5
1. Stoppering	.91	.60	.85	.80
2. Examining		.86	.70	.94
3. Labeling			.91	.89
4. Packaging				.95
5. Filling				

the intercorrelations among ratings on several traits that were made by the forelady of 26 inspector-packers of biological materials. It will be observed that the correlations among the traits are consistently high. This means that the order of individual differences is approximately the same on every trait.

CLASSIFICATION OF RATING METHODS

There is no easy way of classifying all of the various kinds of rating procedures that have been devised. As a working classification the four following types may be distinguished: methods that do not attempt to take into account the size of absolute differences between the individuals rated, methods that provide scales for indicating absolute differences between the rated individuals, check-list methods that provide the rater with a series of scaled behavior descriptions, and methods wherein the rater is forced to choose between two or more different, but apparently equally desirable or equally undesirable, descriptions of behavior.

Ranking Methods. These methods yield a ranking or ordering from best to worst of all individuals comprising the group. The rater simply picks out the individual he considers best, the one he considers next best, etc., and ranks them in order.

Rating-scale Methods. Rating methods that provide some kind of a scale for measuring absolute amounts may be classified into scales of discrete categories and graphic or continuous scales. In the scales of discrete categories, two or more categories are provided, representing different amounts of ability or degrees of the characteristic. The person doing the rating checks the category that he feels best describes the

person being rated. In graphic rating scales, similar categories are provided, but the rater, instead of having to check one or another of the categories, is provided with an uninterrupted line placed just above the category notations, on which he indicates his rating by means of a check mark. The check mark need not correspond exactly with any of the reference points or category notations under the line but may be placed at any point along its length. On some graphic rating scales only the extremes and midmost categories are indicated by means of category descriptions.

Rating scales may also be classified in terms of the kind of standards or categories used. On some scales the rater simply checks a number in indicating the amount of the trait he believes the person being rated possesses. On others he checks one or another of a series of descriptive adjectives or descriptions of graded samples of behavior.

Check-list Methods. These procedures furnish the rater with a check list of prescaled descriptions of behavior. The scale values of the behavior items are unknown to the rater. The task presented to him is to check all of the items that he believes describe the worker being assessed. A final rating is obtained by averaging the scale values of the items he has checked to describe the worker.

Forced-choice Methods. In forced-choice methods the rater is presented with a series of pairs or triads of descriptions of behavior or trait names that are equally desirable (or undesirable). In each pair or triad he chooses the one he thinks most (or least) describes the individual being rated. Each of the alternatives has a scale value which is unknown to the rater. The final rating is obtained by totaling the scale value of the items that have been indicated as describing the individual.

RANKING METHODS

Advantages of Ranking. Rating individuals by arranging them in order of merit has two major advantages: simplicity and naturalness. It is a very simple process to evaluate persons by merely arranging them in order from best to poorest or from most to least in terms of some characteristic. Furthermore, ranking is a very natural type of evaluation, involving a kind of judgment which is frequently made in everyday living. Hence this method is advised for use in situations where the raters are unfamiliar with problems involved in the appraisal of people or are unable to execute the more complicated rating procedures.

Disadvantages of Ranking. There are two major disadvantages of these methods. First of all, the task of ranking a group of individuals becomes difficult when there are over twenty or thirty cases. Secondly, the magnitude of the differences in ability between ranks is not equal at different

positions. For instance, the difference in ability between the fifth and sixth individuals may be much greater, in absolute terms, than the difference between the sixth and seventh. In terms of rank, however, the differences between these individuals are the same. Since absolute differences in ability between individuals are not taken into account, it is apparent that individuals ranked in one group cannot be compared with those ranked in another group.

If only a relatively few individuals are to be rated, and if the only requirement is to learn which ones are better or which ones are poorer regardless of absolute ability, then a simple ranking method may well suffice. With larger groups about whom some notion of absolute ability is important, some other method, such as the rating scale, should be used.

Aids to Ranking. When the number of individuals to be evaluated is large, one of several methods may be used to reduce the labor of ranking. Each rater may simply assign each individual to one of a small number of groups, roughly classifying them as good, average, or poor. Then the individuals within each group are ranked. After possibly a few shifts of borderline cases, the ranked groups are then combined. Similar to this method is the "peeling" process wherein the rater selects the best and the poorest individuals, then he selects the best and the poorest of the remaining cases, and so on, through the entire group. The advantage of this method is that after several extreme pairs have been "peeled" off there are fewer cases to differentiate in the middle range of ability where differences between people are likely to be smaller and harder to distinguish.

Since it is easier to judge which of two workers is superior than to arrange members of a large group of workers in order of excellence, the *method of paired comparisons* sometimes is used. In this method the rater compares each man in the group with every other man. The final ranking of the workers is determined from the number of times each was judged better than the others. Although the ease of judging is probably greater, the amount of work that must be performed is greatly increased. The rater must make $\frac{n(n-1)}{2}$ judgments, where n is the number of men to be ranked. Thus, if there are 50 men in the group, the rater must make 1,225 separate comparisons.

Transmuting Ranks to Values on a Continuous Scale. Ranks are ordinal numbers; *i.e.*, they are numbers that indicate the relative position of an individual with respect to others in a group. The rank of an individual tells *how many other persons* he is better or poorer than, but it does not tell *how much* better or poorer he is than the other members of the group. Because ranks are ordinal numbers, they cannot be arithmetically manipulated as can numbers representing points on a con-

tinuous scale, such as amount of test intelligence, number of items produced per day, or accident rate per year. Consequently, when the workers in a group have been ranked by several supervisors, the final rank of each worker cannot be taken as the arithmetic average of all of the ranks assigned to him unless certain assumptions are made. If some average is required, the median, *i.e.*, the midmost, rank should be used.[11]

FIG. 4-2. Normal frequency distribution curve.

Because of the ordinal nature of ranks and the difficulties involved in handling them arithmetically and statistically, several methods have been suggested for transmuting ranks into values on a continuous scale.[11] In general, these methods are based upon the assumption that any characteristic of human beings that is measured on a quantitative scale will be distributed in accordance with the normal distribution curve. A distribution of this type is shown in Fig. 4-2. If an ability under consideration is normally distributed, then the differences in absolute amount of

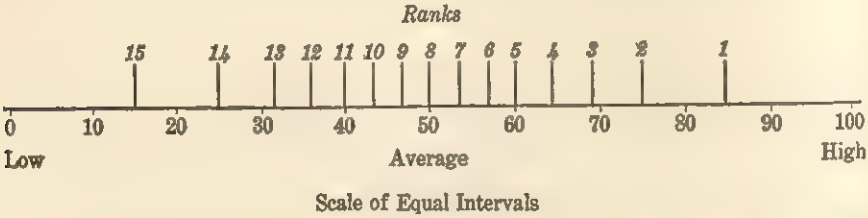


FIG. 4-3. The spacing of 15 individuals along a scale, assuming a normal frequency distribution.

ability between adjacent individuals at either end of the distribution will be greater than the differences in absolute amount of ability between adjacent individuals in the middle portions of the distribution. By way of illustration, the spacing of individuals along a continuous scale, assuming a normal distribution of ability, is shown in Fig. 4-3. The marked differences between ranks at different points on the scale are very apparent.

Table 4-3. Table Based on the Normal Frequency Distribution Curve for Transmuting Per Cent Position in a Ranked Series into Values on a Continuous Scale

Per cent position	Scale value	Per cent position	Scale value	Per cent position	Scale value	Per cent position	Scale value
1	93	26	63	51	50	76	36
2	89	27	62	52	49	77	35
3	86	28	62	53	49	78	35
4	84	29	61	54	48	79	34
5	82	30	60	55	48	80	33
6	80	31	60	56	47	81	33
7	79	32	59	57	47	82	32
8	78	33	59	58	46	83	31
9	76	34	58	59	46	84	31
10	75	35	58	60	45	85	30
11	74	36	57	61	44	86	29
12	73	37	57	62	44	87	28
13	72	38	56	63	43	88	27
14	71	39	56	64	43	89	26
15	70	40	55	65	42	90	25
16	69	41	54	66	42	91	24
17	69	42	54	67	41	92	22
18	68	43	53	68	41	93	21
19	67	44	53	69	40	94	20
20	67	45	52	70	40	95	18
21	66	46	52	71	39	96	16
22	65	47	51	72	38	97	14
23	65	48	51	73	38	98	11
24	64	49	50	74	37	99	7
25	63	50	50	75	37		

On the basis of the mathematical relationship existing between ranks and scores on a continuous scale, assuming a normal distribution, tables have been prepared to facilitate the transmutation of ranks into scaled scores.[12] These scaled scores are so devised that the average is 50, and the extreme values approach 0 and 100. A table of this kind is presented in Table 4-3. In effecting the transmutation the following formula is employed:

$$\% = \frac{100(R - .5)}{N}$$

where the left-hand side is the percentage position, R the rank of the individual under consideration, and N the number of individuals being

ranked. Thus the percentage position of the second-ranking person in a group of 10 would be

$$\% = \frac{100(2 - .5)}{10}$$

$$= 15$$

In Table 4-3 it will be seen that a scale value of 70 is obtained for a percentage position of 15. This value, 70, would be this individual's score

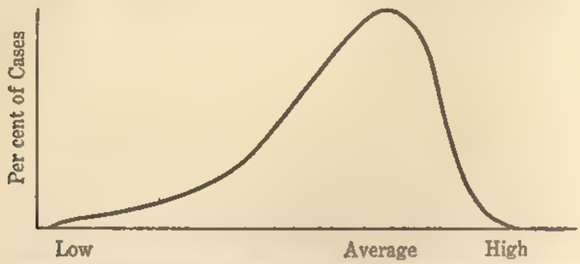


Fig. 4-4. Example of a skewed (non-normal) frequency distribution curve.

on a continuous distribution, assuming that a normal distribution holds within the group of ten individuals for the characteristic on which they were ranked.

Although it is true that many physical and psychological characteristics are found to be distributed normally in the general population, it certainly does not follow that all characteristics in a specially selected group will be distributed in this manner. As a matter of fact, it is un-

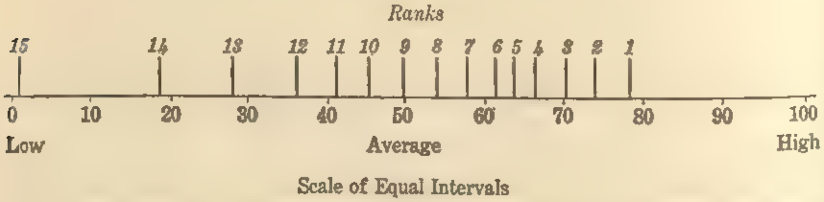


Fig. 4-5. The spacing of 15 individuals along a scale of equal intervals, assuming a distribution skewed as in Fig. 4-4.

likely that a trait such as speed of work would be normally distributed in a highly proficient group of workers. It would be expected that, through a process of selection, elimination, and training, there would be a large number of "high" persons and a small number of "low" persons, a considerable variation from the normal distribution. Such a non-normal distribution is shown in Fig. 4-4. For a distribution of this type the spacing of individuals along a scale would be like that shown in Fig. 4-5.

It is apparent that the differences in absolute amount of ability are greater between adjacent persons at the low end of the scale than between adjacent persons at the high end, and that the smallest differences occur in a region above the average.

In view of the fact that for any group of workers, the shape of the distribution of the trait on which the workers are to be ranked will be unknown, the transformation of ranks into values on a continuous scale will introduce an error of unknown extent. As a consequence, such transformations must be applied with caution.

Group-order Rankings. In some instances a complete ranking of all individuals in order of merit is not necessary. Sometimes what is desired is to know in which half, fifth, seventh, etc., of his group a person belongs. In this case, the group of workers is divided into whatever proportions appear to be desirable. Ratings of this kind are termed group-order rankings and are illustrated in Table 4-4.

Table 4-4. Example of Group-order Ranking Procedure

Highest fifth	Next highest fifth	Middle fifth	Next lowest fifth	Lowest fifth
Outstanding, among best agents in the company	Superior, above average but not outstanding	Average, neither superior nor inferior	Below average but not poor	Falls among the poorest agents in the company

NOTE: For the over-all rating, consider both the individual's success as a casualty insurance agent and his loyalty to the X company. When making this judgment, keep in mind all the X company's agents. The phrases above were chosen to aid in describing successive fifths of all casualty agents in general job success. Therefore, 20 per cent of the agents should be placed under each heading.

When the group-order method is used, it is customary to assign a series of numbers to the categories. If the group is divided into fifths, those in the lowest fifth are assigned a rating of 1, and those in the highest fifth a rating of 5, with the intermediate categories being assigned the appropriate numbers between. These final ratings are then treated as if they were values on a continuous scale. Although this may be satisfactory for rough work, as pointed out earlier, such a procedure in the treatment of ordinal numbers is not logically sound.

To overcome this deficiency the proportion of individuals assigned to each group may be allotted in accordance with the normal distribution curve. By having larger numbers of persons assigned to those groups

in the central positions and fewer to the more extreme groups, an approach to normal distribution can be achieved. If this procedure is accurately carried out, the differences between successive groups in terms of the ability under consideration can be considered to be equal.

In effecting this type of grouping the total range of individual differences is divided into whatever number of categories is desired, and frequencies are set up for each category in accordance with the normal frequency distribution. Table 4-5 gives appropriate proportions for dividing a normal population into varying numbers of steps. Since the effective range of individual differences is greater when the number of persons in the distribution is larger, separate frequencies have been presented for situations where 20, 40, and 60 individuals are to be rated. An example of a rating form employing this principle is shown in Table 4-6.

Table 4-5. Percentage of Cases Expected on the Basis of a Normal Frequency Distribution to Fall into Each Step on Rating Scales of Varying Numbers of Steps

	Number of intervals																				
	3			4			5			6			7			8			9		
No. of cases to be rated	20	40	60	20	40	60	20	40	60	20	40	60	20	40	60	20	40	60	20	40	60
Steps in scale	Per cent																				
1	25	21	20	14	12	10	9	7	5	8	5	5	6	4	4	5	3	2	4	3	2
2	50	58	60	36	38	40	24	25	26	16	16	15	11	11	10	9	8	8	7	6	6
3	25	21	20	36	38	40	34	36	38	27	29	30	20	21	22	15	16	16	12	12	12
4				14	12	10	24	25	26	27	29	30	26	28	28	21	23	24	17	18	19
5							9	7	5	16	16	15	20	21	22	21	23	24	20	22	22
6										8	5	5	11	11	10	15	16	16	17	18	19
7													6	4	4	9	8	8	12	12	12
8																5	3	2	7	6	6
9																			4	3	2

In the earlier discussion concerning the transformation of ranks into values on a continuous scale some of the problems involved in the assumption of a normal frequency distribution were considered. The same limitations hold in the assumption of this type of distribution in the group-order ranking method.

Table 4-6. Rating Scale for Clerical Workers Employing the Group-order Method of Ranking, with Proportions Distributed among the Groups in Accordance with the Normal Frequency Distribution

Rate each of the clerks in your department on the following characteristics. Rate all of them on the first characteristic, Amount of Output, then rate all of them on the second characteristic, Quality of Output, etc. Assign the ratings as follows:

A = 7% B = 25% C = 36% D = 25% E = 7%

These percentages must be adhered to as closely as the number of clerks you are rating will permit. For each characteristic enter all names in the appropriate spaces provided.

	E	D	C	B	A
Amount of Output					
Quality of Output					
Accuracy					
Punctuality					
Initiative					
Interest					

RATING-SCALE METHODS

Types of Rating Scales. As indicated earlier, there are two general types of rating scales: scales of discrete categories and continuous or graphic scales. The standards used are the same for both types of scales. The advantages and disadvantages of the different types of standards will be discussed below, and the conclusions drawn will hold for both types of scales. It is said that the continuous rating scale has some advantage over the scale of discrete categories in that it does not compel the rater to force each individual into one or another of a relatively few categories. This problem is allied to another which is concerned with the number of steps or categories to be used in a discrete scale, a subject that will be discussed later. Examples of several types of rating scales are shown in Table 4-7.

In arriving at a quantitative rating, when scales of discrete categories are used and the categories themselves are not numerical, a series of

Table 4-7. Examples of Rating Scales

Cooperativeness: consider the man's ability to work smoothly and well with others

0 1 2 3 4 5 6 7 8 9 10

Tact in dealing with people

____Excellent
____Good
____Average
____Poor
____Inadequate

Alertness

Very slow in grasping the meaning of things; misunderstands questions	Fairly slow in grasping the meaning of questions or statements	Quick in producing ideas and understanding questions	Exceptionally keen and alert

numbers is arbitrarily assigned to the categories. Thus, in a scale employing four descriptive adjectives such as poor, fair, good, and excellent, a person rated as poor would be given the quantitative rating of zero, a person rated as fair would be given a score of 1, and so on. In another example, if the scale had eleven categories, the lowest category might be assigned the arbitrary value of zero, the next the value of 10, and so on to the highest category, which would be assigned the value of 100.

In continuous or graphic rating scales, a line, which usually is no longer than 5 in. or less than 2, is placed above the categories. If the line is too long the whole scale cannot be adequately grasped by the rater, whereas if the line is too short the rater will find difficulty in accurately estimating distances on it. The rater simply places a check mark directly over a standard or at any point between contiguous standards. In obtaining an individual's rating on the scale, a ruler is placed below the line on which the check mark is made, and the length of the line between the lowest value of the scale and the check mark is read off. The rating may be expressed in terms of the units on the ruler, such as millimeters, in terms of an arbitrary scale, such as from 0 to 10, or in terms of the percentage of the total scale distance occupied by the distance to the lowest value on the scale to the check mark.

Standards Used in Rating Scales. There are four general kinds of standards used in rating scales, namely, numerical or alphabetical, descriptive-adjective, man-to-man, and behavior-sample. The first two classes of standards are simple in nature, and there is little work involved

in constructing them. The last two present several problems, particularly if they are to represent equal steps of amount in the characteristic to be rated.

Numerical and Alphabetical Rating Scales. In numerical rating scales a series of numbers is provided which is supposed to indicate varying amounts of the characteristic being rated. For example, on a scale ranging from 0 to 10, if a person were judged as the best possible worker he would be rated 10. Sometimes 0 is the value assigned to average ability, with a series of positive numbers indicating superior degrees of the characteristic and a series of negative numbers indicating inferior degrees.

The major difficulty with numerical rating scales, and it is indeed a very serious one, is that the numbers which serve as standards are abstract in meaning; they bear no relation to reality. It is not natural to think of an individual as being "6" in quality of output, or "−3" in promptness. The rater is therefore forced to provide for himself some concrete meaning for these otherwise abstract units. This is a source of marked discrepancies in meaning between different raters. One rater may feel that 10 indicates such a high degree of ability that only one person in ten thousand possesses it, whereas another rater may interpret such a rating as indicating a level of superiority that is relatively common.

Similar to numerical scales are those in which letters are used as standards. In a five-step scale the highest category might be *A* and the lowest *E*. There appears to be little choice between using numbers or letters as standards. Any deficiency in the one will certainly be found in the other.

Descriptive-adjective Rating Scales. Scales wherein a series of descriptive adjectives are used overcome to some extent the abstractness that is characteristic of numerical scales. The phrase "very superior" is probably more meaningful than the number "10." Even so, raters will differ in their interpretation of the descriptive adjectives, and thus the standards will not have the same meaning for all raters.

Data presented in Table 4-8 will serve as an illustration of the manner in which raters vary in their interpretation of descriptive adjectives. Fourteen adjectives found in various scales used for rating workers were presented in a randomized list to 34 persons. These persons were told that each adjective described a worker, and on the basis of these adjectives they were to rank the 14 workers in order of merit. The percentages of judges assigning various ranks to each worker are shown. For example, 65 per cent of the group considered the worker who was described as "exceptional" as being the best of the 14, whereas 21 per cent considered the worker described as "superior" as being the best, and 15

Table 4-8. Percentages of 34 Persons Assigning Various Ranks to 14 Descriptive Adjectives When the Adjectives Are Taken as Ratings of Workers

Rating	Rank														Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Exceptional	65	29	6												100
Excellent	15	44	41												100
Superior	21	26	53												100
Good				50	41	9									100
Competent				47	41	6	3		3						100
Satisfactory					6	62	24	3			6				101
Adequate				3	12	12	32	21	15	6					101
Average						6	26	44	12	6	6				100
Ordinary						3	6	15	50	24	3				101
Fair						3	9	12	18	26	29	3			100
Mediocre								6	3	32	44	9	3	3	100
Weak										3	9	71	15	3	101
Poor											3	6	59	32	100
Inadequate										3		12	24	62	101
Total	101	99	100	100	100	101	100	101	101	10[illegible]	100	101	101	100	

per cent considered the worker described as "excellent" as being the best. Although those adjectives obviously indicating superior performance were ranked high and those indicating inferior performance were ranked low, nevertheless there was considerable disagreement in the placement of the words. The worker described as "adequate" was ranked by some individuals as high as 4 and by others as low as 10.

The differences in amount of the trait or ability indicated by standards equally distant apart should be equal; otherwise the scale will not be a scale in the true sense. A measure of linear distance, such as a foot rule, on which the first unit is 1½ in. long, the second ¾ in., the third 2⅜ in., etc., would not be a scale with comparable units. On such a scale an object two units long could not be said to be twice as long as an object only one unit long. Similarly, the difference in ability between two contiguous standards on a rating scale, such as "very superior" and "superior," might not be considered by the rater as being equal to the difference on the same scale between the two next contiguous standards, "superior" and "good." Little or no attempt has been made to find a series of adjectives that a large number of individuals agree indicate equal differences. Indeed, reversal of the order of the same terms is sometimes encountered on different scales. On one scale "superior" might be the step above "average," and "good" the step above "superior." On another scale "superior" might be placed above "good,"

the latter adjective being considered as indicating ability just above "average." *

Man-to-man Rating Scales. In the so-called "man-to-man" rating scales the names of actual individuals known to the rater are used as a series of standards. Originally the procedure in constructing a man-to-man rating scale was to have the prospective rater select 10 to 20 individuals of varying ability whom he knew well and have him rank them in the ability under consideration. The highest-ranking man was used as the highest standard, the lowest-ranking man as the lowest standard, the individual with the midmost ranking as the middle standard, and individuals equally ranked between the middle and the extremes as the remaining standards. In rating another person the rater simply looked over the individuals selected as standards and checked the one nearest in ability to the person being rated.

The foregoing procedure for the construction of the man-to-man rating scale is of course completely incongruous with the basic concept of what constitutes an adequate series of standards for a rating scale, since there is no assurance that the differences between ranks indicate equal differences in ability. The correct procedure is to have the rater select a person of his acquaintance whom he considers average in the ability in question. Two other acquaintances are then selected, the one being as much superior to the "average" person as the other is inferior. The rater then selects a fourth acquaintance who is judged to be halfway between the average and the superior individual, and a fifth who is judged to be halfway between the average and the inferior individual. This procedure is continued until there are obtained as many comparison individuals as the number of standards desired. The procedure is likely to be difficult and laborious even for a rater of considerable experience, but it is necessary if a scale approaching equal intervals is to be achieved.

Besides the difficulty encountered in constructing such a scale there is a further disadvantage in the use of man-to-man rating scales. The standards set up by different raters may not consist of the same individuals and therefore the ratings by different judges will not be comparable. Of course one rater cannot successfully use a man-to-man scale constructed by another.

The obvious advantage of man-to-man rating scales is that the standards are very concrete. They are in terms that even an untrained rater can understand; they are real.

Behavior-sample Rating Scales. Scales employing a graded series of samples of behavior as standards are potentially the most refined of all rating scales. The use of actual descriptions of behavior reduces some

* Procedures for scaling adjectives have been considered by Mosier.[15]

of the differences in interpretation usually found among different raters. For example, on a rating scale measuring the trait "reliability," it is highly probable that the agreement among different raters on the meaning of the sample of behavior "observations and readings of instruments are always checked for accuracy and completeness" will be greater than the agreement on the meaning of the phrase "very reliable" or of the number "5."

In the construction of behavior-sample rating scales, the customary procedure is simply to write down a set of samples of behavior obviously describing persons who vary in the trait to be rated. An individual may be able to construct a satisfactory behavior-sample scale by this procedure if he is very familiar with the problems of ratings and has had considerable experience in the construction and administration of rating procedures. However, the use of a more systematic procedure is of considerable aid in achieving an adequate scale. A procedure is outlined below which greatly increases the chances that the behavior-sample standards will be equally spaced psychologically and will be uniformly agreed upon and understood. The descriptive terms utilized will be in the common parlance of the raters and therefore highly meaningful.

1. On the basis of interviews with supervisors and comments made by them about workers, a number of descriptions of the "best" worker are prepared.
2. These descriptions are submitted to the supervisors with the instruction that they are to pick out the one which most nearly describes the "best" worker. If the extent of agreement among the supervisors is high, that description selected by them is retained as the highest standard on the scale.
3. If there is little agreement among the supervisors on these descriptions, new descriptions are prepared and submitted to them.
4. A similar procedure is followed in developing a description of the "poorest" worker as the lowest standard on the scale.
5. The supervisors are then presented with the two descriptions selected as the extreme standards and are asked to describe a worker exactly halfway between these two, *i.e.*, an average worker.
6. On the basis of these statements, descriptions of an "average" worker are prepared.
7. The best one of these is selected in the manner described above.
8. As many additional standards as appear desirable are developed, using the same procedure as that used for the "average" worker standard.

The Number of Steps in a Scale. A problem that always presents itself in the construction of rating scales involving discrete categories concerns the number of steps or standards to be employed. At present it would seem that no generalizations can be made which will apply to all scales. The best that can be done is to discover empirically the optimal number of steps for each scale. Either the reliability of measurement or the ease of rating can be used as a basis for the determination of the optimal number of steps.

Several factors condition the choice of the number of steps. Obviously, if the steps in a rating scale are very coarse, the raters' powers of discrimination cannot be effectively used. On the other hand, no advantage will be gained if the steps are finer than the ability of the raters to discriminate. In the first case the rating scale will not have maximal reliability, and in the latter the scale will be falsely refined beyond the point of maximal reliability. From a theoretical analysis, Symonds concluded that maximal reliability will be obtained when seven steps are used.[25] However, if the raters are untrained or relatively disinterested, maximal reliability will be reached with fewer steps. The results of an experimental investigation by Champney and Marshall on the relationship between the number of steps in a rating scale and reliability suggest that nine is the optimal number of steps for trained raters.[2] The number of steps yielding the highest reliability, however, probably will vary considerably with the nature of the trait being rated.

There has been much discussion but little actual investigation of the relationship between the number of steps in a rating scale and the ease with which the rater can employ it in the practical situation. The general conclusion is that the fewer the steps the less the difficulty experienced by the rater. Results obtained by Ghiselli, however, suggest that this is not always true.[8] He found that there was greater willingness to express opinion on a four-step scale than on a two-step scale.

Guide Distributions to Encourage the Use of the Entire Scale. When raters fall into the error of leniency or the error of central tendency, the lower part or the extreme parts of the scale are not used. To some extent this is due to the fact that the raters have no knowledge of how individuals should be distributed along a scale of ability. One solution is to provide the raters with a guide, showing the expected proportion of cases that should fall in each category on the rating scale. If the raters are instructed to adhere strictly to these proportions, then a group-order ranking procedure is being employed rather than a rating-scale procedure. In most instances the expected distribution of cases is simply provided as an aid to the rater, and proportions such as those given in Table 4-5 are suggested. Ordinarily a normal frequency distribution is employed. The problems involved in assuming such a distribution have been discussed in connection with the group-order ranking procedure.

Equating Ratings of Different Judges. In the chapter on the measurement of job proficiency, procedures were outlined for achieving comparable measures of proficiency when groups of workers show different levels of productivity because of differences in working situations rather than because of differences in ability. A similar situation exists when several raters assign quite different ratings to the same group of workers.

The differences in the ratings assigned by different raters are to be attributed to their varying susceptibilities to constant errors. Several kinds of mathematical procedures have been devised to equate the ratings assigned by different raters and thereby correct for their idiosyncrasies. For example, regardless of the actual rating values used by the judges, the lowest-rated individual of each judge might arbitrarily be assigned a value of zero and the highest the value of 10. Other persons rated are assigned values on the basis of the proportionality of distance between their rating and the highest or lowest rating.* If such a procedure is followed for the ratings assigned by each rater, then the effects of constant errors of raters would be minimized. Conrad has reported that with raters who have had considerable experience, corrections such as these do not appreciably increase the reliability of the combined ratings.[8]

CHECK-LIST METHODS

Construction of Check-list Scales. Statements representing various levels of worker effectiveness can be accurately scaled in reference to given job characteristics. These statements, in the form of a check list, can then be used by supervisors in evaluating the personnel working under them. The scaling procedure is an adaptation of a method developed by Thurstone for measuring attitude.

Thurstone devised a technique for the measurement of attitude that was based on the older psychophysical method of equal-appearing intervals.[26] This technique yields a series of scaled statements of beliefs or attitudes relative to some particular subject. The items are assigned scale values in terms of the degree to which each is considered a favorable or unfavorable opinion relative to the issue in question. The measurement of an individual's attitude is taken as the average scale value of those items which he checks as representing his opinion.

As was pointed out at the beginning of this chapter, ratings are simply a quantification of opinion. It is apparent that Thurstone's method can be adapted as a means for developing scaled statements for the rating of workers by their supervisors.[18] Following are the steps involved in the preparation of a rating form using the technique of equal-appearing intervals:

1. The job is analyzed in order to discover the major areas in which success should be measured. Some knowledge about the importance of each of these areas also is obtained. This step is necessary, first of all, as a basis for the collection of items, and secondly, as a means for checking the thoroughness of coverage of the job by the final scale.

* A more refined technique would involve the use of standard scores.[11]

2. A large number of statements are collected concerning the performance of individuals on the job. The best source of such statements is the reports made by supervisors on their workers. They may be in the form of comments made by supervisors, opinions expressed on rating scales, or other formal evaluations. Additional statements can be collected through informal conferences with the supervisors, wherein they are asked to describe in specific terms the performance of various workers. Examples of statements from a rating form for salesmen [18] and from a rating form for waitresses are given in Table 4-9.

Table 4-9. Examples of Statements from Check-list Rating Procedures

Examples of statements contained in a rating form for salesmen

Scale value	
29	Is weak on planning
32	Is somewhat in a rut on some of his brand talks
46	Is a good steady worker
56	Tends to keep comfortably ahead on his work schedule
69	Is making exceptional progress

Examples of statements contained in a rating form for waitresses

Scale value	
13	Slow and clumsy
22	Is impatient while taking orders
26	Gets rattled when rushed
36	Does no more than is necessary
39	Works slowly and methodically
57	Neat in appearance and work
65	Cheerful disposition, doesn't get annoyed at customers

3. The statements are edited to ensure that they are specific and to the point. They should not be double-barreled in meaning but should be so formulated that they can either be accepted or rejected. They should be related to successful and unsuccessful performance on the job as indicated by the job analysis. Statements are rejected if they fail to meet these criteria.

4. A number of judges are asked to evaluate each of the edited statements on a scale relative to the degree of job success a worker so described would have. Ordinarily a numerical rating scale with any number of steps from five to eleven is provided for the judges, and every judge rates each statement on the scale. The items in the scales for salesmen and for waitresses in Table 4-9 were evaluated on a scale of seven steps, the lowest being given the value of 10 and the highest, 70.

The scale value of each item is obtained by taking the average of all the judges' ratings. In using the final scale, the scale values are not reproduced on the form and are unknown to the rater.

Apparently any type of group can be used as judges in the evaluation of the statements as long as the individuals are mature and capable of making accurate judgments about job characteristics. In one study such differing groups as foremen, college students, technical assistants, and college faculty members evaluated over five hundred items.[28] The scale values assigned to any given statement by the several

groups were quite similar. In fact the correlations between the scale values assigned by the different groups were exceedingly high, averaging about .98.

5. Those items on which the judges show wide disagreement in their ratings are discarded. Obviously they are ambiguous and have different meanings for the various judges. Only items are retained on which there is a high degree of agreement. In Table 4-10 are given the judgments on two statements of 25 individuals who were

Table 4-10. Examples of Ambiguous and Unambiguous Items on Rating of Waitresses Where Degree of Ambiguity Is Measured in Terms of the Agreement among Ratings of Judges

Item	Scale value							Total no. of judges
	10	20	30	40	50	60	70	
Fast but makes mistakes in orders	3	3	10	7	2			25
Is impatient while taking orders	1	18	6					25

asked to scale a number of statements dealing with the job of waitress. For the first item, "Fast, but makes mistakes in orders," there was low agreement among the judges as to the type of worker it described. Three judges rated the item 10, indicating that they believed it described the poorest type of waitress. Two other judges rated it 50, indicating that they believed it described one who was slightly better than average. On the other hand, with the statement, "Is impatient while taking orders," there was a high level of agreement. All judges agreed that it described an inferior waitress, with more than two-thirds of the judges assigning the item a value of 20.

6. As a final step, statements are selected that represent all levels of ability on the scale and which also cover all aspects of the job as indicated by the job analysis. Depending upon the total number of items desired in the scale, one or more statements are selected that represent proficiency at levels varying by approximately equal amounts from the low to the high end of the scale.

In using a check list of scaled statements the supervisor simply checks those items on the form which he feels describe the particular worker being evaluated. No restrictions are put on the supervisor as to the number of items he should check. Experience indicates that usually a satisfactory number will be checked. The worker's final rating is taken as the average of the scale values of all items that his supervisor has checked in describing him.

Evaluation of General versus Special Areas of Job Performance. From the foregoing description of the manner in which scaled check lists are constructed it is apparent that the ratings yield an appraisal of the over-all worthiness of the worker. Indeed, one of the criticisms of this method is that it cannot be used to indicate areas of strengths and weak-

nesses as is possible with ranking and rating methods where separate evaluations can be made of several different traits or characteristics.

There is no reason; however, why scaled check lists cannot be constructed to measure several different characteristics of worker performance. Statements collected concerning job performance could be classified in terms of areas such as productiveness, cooperativeness, and leadership. Items within each area could be scaled separately so that a rating within each area could be obtained. In order to reduce halo error, statements within each area could be chosen not only in terms of the criteria described above, but also in terms of having low correlations with statements in all of the other areas.

Evaluation of Scaled Check-list Rating Methods. The fact that the rater does not know the exact scale values of the items may lead to the false presumption that with this method rater biases cannot operate. But even a most obtuse rater would know that he would be giving a person a low rating if he checked the item "Does as little as possible" and would be giving him a high rating if he checked the item "Is very resourceful." This method, then, does not overcome rater bias any more than do ranking and rating scale methods. Its advantage lies in the fact that the procedure permits the rater to make more precise and less ambiguous expressions of his opinion concerning the worthiness of the individual being judged.

Although no studies have been made concerning the extent of constant and halo errors in this type of rating as compared with other rating methods, in view of the specific nature of the behavior described in the statements it would appear likely that this procedure would tend to minimize these errors.

As far as reliability is concerned, striking results have been reported. Richardson and Kuder found that the coefficient of correlation between ratings of 305 salesmen made by two independent raters on a 36-item form was .83.[18] Knauft developed a rating form of scaled items for the appraisal of bakeshop managers on which he obtained a correlation of .81 between independent ratings made by two supervisors.[14] These findings indicate that check lists of scaled items give high reliability when developed for specific jobs.

In some situations it is desirable to obtain a rating on over-all performance for comparing persons working on different jobs. Goertzel has developed a rating form of scaled items for such generalized use.[10] Application of this form to various job groups gave reliability coefficients as high as those reported above for specific jobs. These findings suggest that such generalized rating forms will give accurate appraisals of workers on jobs differing in many characteristics. Uhrbrock has published over

seven hundred scaled statements of this type that can be utilized for generalized rating forms.[28]

FORCED-CHOICE METHODS

The one general criticism that is made of the rating methods discussed thus far is that the evaluations they yield are readily subject to any biases that the raters might have. With these methods distortions of the facts can readily be introduced. If a supervisor wishes to discriminate against a particular worker he simply can disregard the instructions on the rating form and rate the individual low. Again, a supervisor may make his appraisals in a purely superficial fashion, rating all individuals at about the same level. It is apparent that there are two possible ways of reducing these errors. One is to train the raters in rating procedures and to motivate them to make accurate appraisals, thus establishing attitudes in them that will minimize their personal biases. The other is to develop a rating procedure in which personal biases cannot influence the ratings. The forced-choice rating methods are attempts to provide a solution of the second kind.

Forms of Forced-choice Ratings. The forced-choice rating form appears as a series of groups of statements or items, and the rater chooses among the members of each group in terms of how well he believes the statements describe the individual being rated. Various arrangements of items have been employed. Sometimes the items appear in pairs or triads of favorable descriptions as shown in the top part of Table 4-11. In this case the rater checks the one item in each group which he believes best characterizes the individual being rated. In some instances all of the descriptions in a group are unfavorable. Raters tend to feel that they are being put in an untenable position if they have to indicate one of several undesirable characteristics as describing an individual. Therefore the usual practice when such items are used is to have the rater check the item that is least descriptive. In all cases the rater is forced to choose among items that are equally desirable or equally undesirable.

In another form of the forced-choice method each of the groups contains four items, two descriptions of desirable characteristics and two of undesirable characteristics. Examples of such groups of items are given in the lower part of Table 4-11. In this form the rater makes two checks in each group, one for the item which best describes the individual and one for the item which is least descriptive. In some instances trait names such as honest, dynamic, accurate and businesslike are utilized rather than behavior descriptions, but the forms of the rating schedule are the same as shown in Table 4-11.

Table 4-11. Examples of Items Contained in Forced-choice Rating Forms

In each group place a check mark in front of the statement that you believe best characterizes the employee under consideration. Be sure to put only one check mark in each group of statements.

______Is very patient
______Arrives at conclusions logically
______Assumes responsibility for his own mistakes

______Delegates work very wisely
______Is exceptionally fair
______Inspires his associates

In each group place a check mark in front of the statement that you believe best characterizes the employee under consideration, and another check mark in front of the statement that you believe least characterizes him. Be sure to put two and only two check marks in each group of statements.

______Has a well-rounded personality
______Lacks force and drive
______Tends to be overbearing
______Shows foresight

______Displays disloyalty
______Is almost indispensable
______Makes many mistakes
______Has a very promising future

Scoring the Statements. Each of the statements has a previously determined weight. As in the case of check-list scales these values are unknown to the rater. An individual's rating is determined by adding together the values of each of the statements that have been checked as describing him. Suppose the following items had the weights indicated in parentheses and the rater checked the one member of each pair that he thought best described the individual.

(0) Honest
(3) Careful

(1) Hard working
(0) Cooperative

(2) Shows foresight
(0) A well-rounded personality

An individual for whom the alternatives "Careful," "Hard working," and "Shows foresight" were checked would have a rating of six. One for whom the alternatives "Honest," "Hard working," and "A well-rounded personality" were checked would have a rating of one. In any final scale, of course, there would be far more than just three groups of items. The

larger the number of groups the more reliable the scale is expected to be. In practice, scales have been used with as few as 10 groups of items and with as many as 50.

Construction of Forced-choice Scales. Procedures for the construction of forced-choice scales have been described by several authors.[19, 21, 23] The first three steps, which are concerned principally with the collection and editing of items, are the same as the first three steps followed in the construction of check-list scales. Beyond these steps, however, the procedures differ considerably.

There are two bases on which an item is selected for inclusion in a forced-choice scale, its *preference value* and its *discriminative value.* Preference value refers to the degree to which an item is considered to reflect a desirable or undesirable characteristic. Discriminative power reflects the degree to which the item distinguishes between "good" and "poor" workers. Since the alternatives of an item have the same preference value, that is, they are equally desirable or equally undesirable, the rater's personal biases presumably cannot operate in the selection of the statements he checks for the individual being rated.

Determination of the Preference Value of Alternatives. Preference value has been determined in several different ways. The method which would seem to be most meaningful is to have a group of persons evaluate each statement in terms of its desirability for them. For example, judges might be instructed to rate each statement in terms of how well they themselves would like to be described by it, particularly if they were employed on the job in question. Ratings of each statement might be made on a scale such as the following:

6. Like extremely
5. Like strongly
4. Like mildly
3. Indifferent to
2. Dislike mildly
1. Dislike strongly
0. Dislike extremely

Another procedure is to have the raters evaluate each of the statements in terms of general desirability on a scale like the following:

5. An exceedingly desirable characteristic
4. A very desirable characteristic
3. A fairly desirable characteristic
2. A fairly undesirable characteristic
1. A very undesirable characteristic
0. An exceedingly undesirable characteristic

In this procedure the judgment of the rater is less personal and more general and abstract.[*]

The first of the above methods would seem to yield the more meaningful indices of preference value. The ultimate aim is to put into the same grouping descriptions of behaviors or traits which connote to the raters equally desirable or equally undesirable characteristics. This probably is best achieved by having the judges use themselves as the frame of reference in determining whether a particular behavior is desirable or not.

It would seem desirable to have the preference rating made by individuals similar to those who ultimately will use the scale. If line foremen are to use the scale then line foremen should make the preference judgments. In the case of check lists the results indicate that it makes little difference what type of group evaluates the statements. Preference ratings, however, involve personal values rather than evaluations of goodness of performance. Thus, for example, college students may not care one way or another about being described as businesslike, whereas it would be a very important characteristic in the eyes of office workers.

Sometimes preference values are obtained in a different way. Each of a number of raters is asked to consider some individual with whom he is acquainted and who is employed on the job for which the rating form is being constructed. The raters are presented with statements and are asked to judge the degree to which each statement characterizes the individual under consideration in a manner such as the following:

4. To an exceedingly high degree
3. To quite a high degree
2. To a fair degree
1. To a slight degree
0. Not at all

The preference value of each statement is taken as the average of these ratings. Those statements with high average ratings are considered to have high preference value and those with low ratings low preference value. It is apparent that ratings of this kind have a very different meaning than the ratings of preference value discussed above. It is not quite clear exactly what meaning is to be ascribed to these ratings unless they are interpreted as being indices of popularity of the statements. In any event it can be said that alternatives with equal ratings determined by this procedure still may differ in terms of their desirability as judged by different raters. It is one thing to say that two different traits characterize a particular person equally well, and quite another to say that these two traits are equally desirable. It can therefore be concluded that these values based upon popularity of use should not be utilized as preference values.

The fourth step in the construction of forced-choice scales, then, is to have each of the statements rated in terms of preference value. The index

of preference value for a statement is determined by averaging the ratings assigned to it by a number of judges.

Discarding Ambiguous Statements. Although to the writers' knowledge it has never been applied in evaluating the preference value of statements for forced-choice rating scales, the procedure outlined in step 5 in the development of check lists certainly should be accomplished. This procedure is concerned with weeding out statements that are ambiguous, that is, statements placed widely apart on a preference scale by different raters. It is quite apparent that some people may think a particular characteristic is highly desirable while others may have the opposite view. For example, some may feel that they are describing an employee favorably by characterizing him as "aggressive," while for others describing a person in this manner is tantamount to damning him. The fifth step, then, is to evaluate statements in terms of ambiguity and to discard those which are too ambiguous.

Determination of Discriminative Value of Statements. The sixth step in the construction of a forced-choice scale is to evaluate each of the statements in terms of discriminative value. It will be recalled that discriminative value refers to the degree to which a statement distinguishes between good and poor workers. Obviously, therefore, it is necessary to have a measure completely separate from preference value, in terms of which goodness of performance can be gauged. By way of example, suppose it were desired to construct a forced-choice rating form that could be used by supervisors for rating foremen. In determining discriminative values the first step would be to present the statements to the supervisors in a list within which the statements were randomized with respect to preference value. The supervisor would check every statement in the list which he felt was descriptive of each foreman. Then the foremen would be divided into "good" and "poor" groups in terms of some index of proficiency on the job. In the present example this might be made on the basis of whether the foreman's department was high or low in production. Finally the per cent of good and poor foremen for whom each statement was checked would be obtained and the weight of each item determined. Illustrations of the necessary calculations are given in Table 4-12. The first statement in this table has a low discriminative value because it hardly discriminates at all between good and poor foremen. The remaining statements show varying degrees of discriminative value.

Another procedure for determining discriminative value is to have the supervisors check the statements which they consider characterize a given foreman, and then finally rate him on an ordinary rating scale. The division of foremen into "good" and "poor" for purposes of obtaining discriminative values is made on the basis of these preliminary ratings. Discriminative values could also be taken as the difference between the

Table 4-12. Calculations for Obtaining the Weights for the Discriminative Values of Statements in a Forced-choice Rating Form for Foremen

% of foremen for whom item is checked		Difference	Discriminative weight	Item
Foremen in high production group	Foremen in low production group			
75	74	1	0	Makes many useful suggestions
81	68	13	1	Thoroughly understands his own weaknesses
92	72	20	2	Handles people very well
85	53	32	3	Clearly imparts information to others

average preliminary rating of the foremen for whom the item is checked and the average preliminary rating of those for whom the item is not checked.

When preliminary ratings are utilized as a basis for determining discriminative values of statements it is necessary that these ratings be both reliable and free from bias. However, it would appear that there would be a contamination between checking off statements and making preliminary ratings. This circumstance would result in a false validity. Raters who are unfairly biased against certain workers not only would rate them low in the preliminary ratings but also might tend to check one rather than another of two alternatives for these workers on a forced-choice item. If satisfactory appraisals can be obtained by ordinary rating methods then the question can be asked why a forced-choice procedure is necessary. It might be, of course, that raters would be willing to make honest and accurate ratings for research purposes in the development of a forced-choice procedure, whereas under ordinary operating conditions they would not. According to this argument the presumption is made that those who ultimately will use the forced-choice rating form are impelled to make inaccurate ratings. Therefore, it is assumed that a forced-choice procedure wherein the discriminative values of statements are developed on the basis of ratings by honest and dependable raters will give accurate results in the hands of raters who are inexperienced or are poorly motivated. At the present time there is no evidence on which the validity of this assumption can be assessed.

Forming Statements into Item Groups. The seventh and final step is to form groups of statements into pairs, triads, etc., in which the statements will have the same or quite similar preference values but quite different discriminative values. The weights for the alternatives in any given item are a function of the differences in the discriminative values of the alternatives. Thus if the first and last statements in Table 4-12 were put into a forced-choice pair, the first would be assigned a weight of zero and the last a weight of 3. The first statement is equally characteristic of good and poor foremen while the last is much more characteristic of good than of poor foremen.

Evaluation of the Forced-choice Rating Method. The forced-choice rating method is so very new that there have not been a sufficient number of evaluative studies to warrant any final conclusions concerning its usefulness. Certainly it is premature to conclude, as is so often done, that the method provides the solution to all of the difficulties inherent in the older rating methods.[27] While the reliability of ratings obtained by forced-choice methods has been reported to be substantial, there have been too few pertinent investigations to justify precise generalizations. Similarly, studies of constant and halo errors are lacking.

Because the method results in the grouping of statements which are equal in terms of social desirability, it has been suggested that raters will find it very difficult to make choices among the statements in a group, and therefore will react negatively to the rating procedure as a whole. With such an attitude the raters might then check the statements in a purely superficial way. A study was made in which a rating schedule, for which discriminative indices were already determined, was used with a second group and the discriminative indices checked. The discriminative power of the scale was shown to hold up in the second study.[21] It would seem, then, that this problem of the raters finding the judgments extremely difficult is not a particularly serious one if the raters try to express honest reactions to the statements.

From one point of view it could be said that the purpose of the forced-choice procedure is to "trick" the rater into a situation wherein his personal biases cannot operate. This is accomplished by compelling him to choose among equally desirable or equally undesirable alternatives. If this argument is true then it follows that the raters are persons who either cannot be trusted to make impartial evaluations or are operating in a social situation which forces them away from impartiality. It would appear that if the latter situation exists it would be appropriate for management to take remedial steps to correct the entire situation. It is likely that such an unhealthy atmosphere would show effects in more than just unreliable ratings. It might be expected that under these conditions high labor turnover, low production, and poor morale would be

found. The solution to such a problem is a direct attack upon the entire unhealthy situation and not the development of a new and different rating procedure.

Clearly the development of a forced-choice rating procedure requires more time and is more costly than the development of other types of rating procedures. Particularly in the case of rankings and rating scales, a procedure can be devised in a very short time. The forced-choice method necessitates a considerable amount of time for the collection and evaluation of statements, and could not readily be installed and utilized.

The forced-choice method has promise, and warrants further study. As compared with the older rating procedures it necessitates a more exact consideration of what is being rated by requiring the determination of the discriminative power of the statements. If the discrimination values are based upon poor preliminary ratings then they are of doubtful significance. When they are based upon some independent and objective measure of job success they should be dependable. The question arises as to why these objective measures themselves are not used to evaluate the personnel rather than using forced-choice ratings. The one case where this question is not pertinent is when the objective measures are not immediately available. For instance, the persons who are divided into good workers and poor workers for purposes of computing the discriminative values of the statements might be so classified on the basis of whether they continue on the job or leave the employ of the company. In such an event, of course, determination of the discriminative values could only be accomplished sometime after the individuals were evaluated on the criterion of remaining with or leaving the company.

POOLED INDEPENDENT RATINGS COMPARED WITH JURY RATINGS

In those situations where each worker is evaluated by more than one rater, it is commonly believed that if the raters can discuss each case among themselves, their combined jury rating will be more accurate than a simple pooling or averaging of their individual ratings made independently of one another. Jury ratings by being a group decision would seem to permit of greater accuracy than a simple pooling of individual judgments. Although this problem is of considerable importance, it has not received thorough experimental study.

Rusmore attempted to study the effectiveness of jury ratings as compared with pooled independent ratings in a situation closely paralleling the industrial situation.[20] He had 86 persons rate the skill of 12 operators performing an industrial task. Although the raters were inexperienced, they had had some instruction in rating procedures. Each rater made his

judgment independently. Then after discussion in small groups another rating representing the consensus of the group was obtained. The accuracy of the ratings was determined by correlating them with actual measurements of the speed and accuracy of performance of the workers who were observed. The coefficient of correlation between the combined independent ratings and actual performance was .84. The similar coefficient for the jury ratings was .77. The results of this investigation indicate that jury ratings are not superior to combined independent ratings and even may be slightly inferior.

These findings are contrary to expectation and hence are in need of some explanation. When raters work together in accomplishing their ratings certain social interactions will occur among the group members. It might be anticipated that the opinions of the designated leader or of the individual who informally assumes leadership would carry greatest weight. As the other members of the group subordinate their opinions to his, there will in effect be a reduction in the number of raters. The ratings then will be less reliable since they will not be based upon the joint opinions of the entire group but rather principally upon the opinions of the leader.

The type of situation described above is likely to occur when supervisors submit their ratings to the review of higher authority. In making their ratings they might be influenced by the way they believe their superiors will react. The remedy is to introduce more democratic procedures.

RATINGS ON SPECIFIC ASPECTS OF JOB PERFORMANCE

When rating scales first came into popular use as instruments for measuring the effectiveness of human performance, the usual method was to rate each individual on the single characteristic of general worthiness. In seeking to improve the dependability of these ratings, efforts were directed toward improving the physical aspects of the scale, increasing the definition of the steps, determining the optimal number of steps, etc. More recently the tendency has been to analyze the job into its important components and to prepare a separate scale for each component. An individual's final rating is taken as the average or sum of the ratings on these specific aspects of the work. In some cases the separate scales are given different weights in the determination of the final composite ratings.

Ratings on Specific Aspects of Performance versus Ratings on Over-all Performance. Before the decision is made to base the final appraisal on the average of the ratings from several component scales rather than on a single scale of over-all worthiness, it would be well to compare the two

procedures. Unfortunately the evidence is fragmentary and final conclusions cannot yet be drawn.

Several investigations conducted by the U.S. Employment Service are relevant to the problem.[24] In one study a group of department-store salespersons were rated by two superiors on ten specific aspects of job performance, such as knowledge of merchandise, sales talk, ability to satisfy customers' needs, etc. The coefficients of correlation between the two judges' ratings on the ten specific scales ranged from −.05 to .28, with a mean of .14. When the ten ratings by each judge for each salesperson were averaged, the coefficient of correlation between the average ratings of the two judges was found to be .32. In addition to these specific ratings, the two superiors rated each person on a rating scale of over-all job success. The coefficient of correlation between the over-all ratings of the two judges was .58.

In another investigation by the U.S. Employment Service, a group of cafeteria-counter workers was rated by three judges on two separate occasions. On both occasions the employees were rated on 12 different scales measuring specific aspects of job performance, such as ability to suggest items to customers, economy in handling foods, speed of movement, etc. A rating of over-all general value was also made. The coefficients of correlation between the ratings of pairs of judges on the specific aspects of the work ranged from −.09 to .56, with an average of .32. The coefficients of correlation between the over-all ratings of pairs of judges ranged from .54 to .56, with an average of .55. These last coefficients do not differ greatly from the reliabilities obtained for the composites of the specific ratings which averaged .61 and ranged from .57 to .67.

When the ratings from the two occasions were compared the results showed the same trends. In this case each judge's ratings on the first occasion were correlated with his ratings on the second occasion. The coefficients of correlation between the two sets of ratings for the scales of specific aspects of job performance ranged from .17 to .96, with an average of .52. The coefficients between the composite of these specific ratings ranged from .82 to .92, with an average of .84. The coefficients between the ratings of over-all effectiveness ranged from .77 to .83, with an average of .80.

Bingham has suggested that in order to have adequate reliability, ratings of over-all effectiveness must be made after the rater has first rated on scales of specific aspects of the job.[1] This suggestion is made because if a rater simply makes a general judgment of an employee's value to an organization without first analyzing his specific defects and superiorities, the judgment of over-all effectiveness will be undependable. To examine the validity of Bingham's argument would require a

situation where one judge rated the employees on specific aspects of the work and another made ratings of the over-all performance. If the ratings of the two judges were done independently, one type of rating could not be said to affect the other. Ghiselli investigated just such a situation with the inspector-packers in a pharmaceutical manufacturing plant.[9] In this case the job was broken down into five specific aspects of the work, such as inspecting, labeling, stoppering bottles, etc. The fore-lady rated every girl on scales measuring each of these important aspects of the work. The supervisor rated the girls on a scale of over-all general value to the organization. The coefficient of correlation between the average ratings of the forelady and the ratings of the supervisor was found to be .73. This coefficient indicates that the two types of ratings, although made independently, are fairly similar. More evidence is necessary on this point, however, before deciding on the merits of the two procedures.

From the investigations cited several tentative conclusions can be drawn. First of all it seems apparent that the reliability of ratings of specific aspects of job performance is likely not to be wholly satisfactory. Secondly, the reliability of ratings of specific characteristics is lower than the reliability of both ratings of over-all effectiveness and composite ratings formed from several scales measuring specific characteristics. Thirdly, the reliability of a composite of specific scales is about the same as that of a rating of over-all effectiveness when all judgments are made by the same raters. Lastly, ratings of over-all worthiness seem to be fairly similar to composite ratings based on measures of several specific worker characteristics.

The Weighting of Ratings of Different Aspects of Job Performance. In the chapter on the measurement of job proficiency it was pointed out that when success on a job is evaluated in several ways, the problem arises of optimally weighting each of the measures before combining them into a composite measure of proficiency. A similar problem arises when workers are rated on a number of different aspects of their performance on the job, as, for example, when a form such as that shown in Table 4-6 is used. Ordinarily, the weights assigned to the various scales are based on the judgments of persons familiar with the job. The various methods for combining measures of proficiency discussed earlier can also be employed in determining the best possible way for combining several ratings.

One important consideration is whether the differential weighting of the various scales on a rating form improves the results. If the order of individual differences among the workers in the final ratings is the same when the separate scales are equally weighted as when differential weights are applied, then it can be said that the differential weighting

adds no improvement to the measurement. The coefficient of correlation between the final ratings when the items are equally weighted and the final ratings when they are differentially weighted provides the necessary index. If the coefficient is very high, then the weighting system adds nothing, whereas if the coefficient is moderate or low, the weighting system can be said to be contributing.

In most instances where several rating scales are used the conditions are such that the application of differential weights to the different scales will have little effect. It can be shown mathematically that the correlation coefficient between the final ratings based on equal weights and the final ratings based on differential weights becomes higher under the following conditions: the greater the number of traits being rated, the higher the intercorrelations among ratings of the different traits, and the more similar the weights given different traits (*e.g.*, weights of 4, 5, 6, and 7 as compared with weights of 1, 4, 7, and 10). With a set of only four scales, with intercorrelations among them of approximately .50 (a figure typical for ratings), and weights as disparate as 1, 4, 7, and 10, the coefficient of correlation between the final composite based on differential weights and the final composite based on equal weights will be .97. Since most rating forms employ more than four scales, have intercorrelations among ratings which usually run .50 or higher, and use weights which are usually less disparate than those given in the example, it can be concluded that differential weighting will not give better results than equal weighting of specific scales.

The amount of variation of the ratings on the different scales contributes to the weighting of the scales. Ordinarily it is found that on a rating form containing a number of different scales the ratings on some scales cover many more steps than the ratings on other scales. In the final composite rating those scales with the greatest variation in ratings carry the most weight. This can be shown very simply. Suppose a group of workers was rated on two characteristics, speed of production and quality of output, each set of ratings being made on a scale ranging from 1 to 5. Suppose that on speed of production some workers were given ratings as low as 1 and others ratings as high as 5, while on quality of output all workers were given the rating of 4. When the final rating is computed for each worker by adding his ratings on the two scales, it can be seen that the rating on quality of output will have no differential effect upon the composite ratings, since it is simply a constant number added to each worker's rating on speed of production. It is apparent then that when only a very narrow region of a scale is used by a rater, the ratings on that scale are contributing very little toward the differentiating of the workers being evaluated.

REFERENCES

1. Bingham, W. V.: Halo, invalid and valid, *J. Appl. Psychol.*, **23**, 221–228, 1939.
2. Champney, H., and H. Marshall: Optimal refinement of the rating scale, *J. Appl. Psychol.*, **23**, 323–331, 1939.
3. Conrad, H. S.: The bogey of the "personal equation" in ratings of intelligence, *J. Educ. Psychol.*, **41**, 267–293, 1932.
4. Ferguson, L. W.: The value of acquaintance ratings in criterion research, *Personnel Psychol.*, **2**, 93–102, 1949.
5. Ferguson, L. W.: The effect upon appraisal scores of individual differences in the ability of superiors to appraise subordinates, *Personnel Psychol.*, **2**, 377–382, 1949.
6. Freyd, M.: The graphic rating scale, *J. Educ. Psychol.*, **14**, 83–102, 1923.
7. Furfey, P. H.: An approved rating scale technique, *J. Educ. Psychol.*, **17**, 45–48, 1926.
8. Ghiselli, E. E.: All or none versus graded response questionnaires, *J. Appl. Psychol.*, **23**, 405–413, 1939.
9. Ghiselli, E. E.: Tests for the selection of inspector-packers, *J. Appl. Psychol.*, **26**, 468–476, 1942.
10. Goertzel, V.: An objective rating form for job success, M.A. Thesis, University of California, 1941.
11. Guilford, J. P.: "Psychometric Methods," McGraw-Hill, 1936.
12. Hull, C. L.: "Aptitude Testing," World, 1928.
13. Kingsbury, F. A.: Analyzing ratings and raters, *J. Personnel Research*, **1**, 377–383, 1922.
14. Knauft, E. B.: Construction and use of weighted check-list rating scales for two industrial situations, *J. Appl. Psychol.*, **32**, 63–70, 1948.
15. Mosier, C. I.: A psychometric study of meaning, *J. Soc. Psychol.*, **13**, 123–140, 1941.
16. Paterson, D. G.: Methods of rating human qualities, *Ann. Acad. Political and Social Sci.*, **110**, 81–93, 1923.
17. Remmers, H. H., and M. J. Plice: Reliability of ratings at Purdue University, *Ind. Psychol.*, **1**, 717–721, 1926.
18. Richardson, M. W., and G. F. Kuder: Making a rating scale that measures, *Personnel J.*, **12**, 36–40, 1933.
19. Richardson, M. W.: Forced-choice performance reports; a modern merit-rating method, *Personnel*, **26**, 205–210, 1949.
20. Rusmore, J. T.: An experimental comparison of the composite and jury methods of obtaining group judgments, Ph.D. Thesis, University of California, 1944.
21. Sisson, E. D.: Forced choice—the new Army rating, *Personnel Psychol.*, **1**, 365–381, 1948.
22. Slawson, J.: The reliability of judgments of personal traits, *J. Appl. Psychol.*, **6**, 161–171, 1922.
23. Staff, Personnel Research Section, A.G.O.: The forced choice technique and rating scales, *Am. Psychol.*, **1**, 267, 1946.
24. Stead, W. H., and C. L. Shartle: "Occupational Counseling Techniques," American Book, 1940.
25. Symonds, P. M.: "Diagnosing Personality and Conduct," Appleton-Century-Crofts, 1931.

26. Thurstone, L. L.: Attitudes can be measured, *Am. J. Sociol.*, **33**, 529–554, 1928.
27. Travers, R. M. W.: A critical review of the validity and rationale of the forced-choice technique, *Psychol. Bull.*, **48**, 62–70, 1951.
28. Uhrbrock, R. S.: Standardization of 724 rating scale statements, *Personnel Psychol.*, **3**, 285–316, 1950.

CHAPTER 5

Principles and Problems in the Selection and Classification of Workers

The need for evaluating the potentialities of men is as great in business and industry as in any field of human activity. An individual's realization of success in his work with the consequent personal happiness and satisfaction it brings to him, together with the nature and amount of his contribution to the progress and profits of the organization for which he works, are directly conditioned upon matching the requirements of the work with the individual's abilities. To accomplish this, there is need for developing measuring devices which will permit accurate appraisals of the individual's capacities to function successfully in different work situations.

On any job, workers vary in the quality of their performances, and in many instances the extent of these individual differences reaches large magnitudes. In extreme cases the output of the best workers may be ten to twenty times greater than that of the poorest. Similarly, some workers achieve a wealth of satisfaction from their jobs while others heartily dislike everything connected with theirs. If any significance is attached to these differences in behavior among workers on the job the forecasting of occupational success assumes considerable importance. While differences in job performance among workers can be reduced and certain deficiencies overcome by training, standardization of work methods, and similar procedures, the problem of improving job performance and satisfaction is greatly facilitated if steps are taken at the very beginning to match the capacities of the worker with the nature and requirements of the job he is expected to perform.

THE GENERAL PROBLEMS OF PERSONNEL PLACEMENT

The Problem of Matching Workers and Jobs. When all of the various problems that arise in connection with the placement of workers on jobs are reviewed, it becomes apparent that there are many situations that

should be considered. Most of these can be subsumed under the general problems of personnel selection and personnel classification. In both problems there are, on the one hand, vacant jobs, and on the other, people who can fill these jobs. The task is to effect the best matching of jobs and individuals. The success of the matching is ultimately measured in terms of the adequacy with which previously established objectives are achieved. These objectives may be higher production, greater job satisfaction, lower turnover, etc., or some combination of such measures. If the matching of workers and jobs is successful then the fulfillment of these objectives will be closely approached.

The Problem of Selection. In personnel selection the concern is with the filling of one or more vacancies occurring in a given job. For example, it may be that an organization has openings for three typist-clerks. The task is to decide which, if any, of the available applicants are suited for the job. Essentially the task is one of evaluating, in terms of the requirements of the job, the capacities of those persons who are candidates for it. The problem, therefore, consists in accepting or rejecting workers with respect to a given job.

The Problem of Classification. In classification the situation is just the opposite. It consists in accepting or rejecting various unfilled jobs with respect to an individual who is unplaced or who is potentially a candidate for these openings. In classification there is no concern with accepting or rejecting the individual. The task is to place each worker in that job where he can make the greatest contribution to the organization and to himself. A good example of the classification situation exists in the armed services. At the beginning, to be sure, there is some rejection of men deemed inadequate for military service. Beyond this point, however, the problem is to place every individual in that military occupational specialty where he will be of greatest benefit to the service. Similar situations arise in industry. For example, an organization may have many different kinds of job openings. The problem is to assign every person who has passed an initial screening for over-all minimal acceptance to one or another of the vacant jobs. In classification, then, every person having the minimum qualifications is placed in some kind of job, and preferably in that job in which he can contribute the most to his organization and to himself.

PROBLEMS IN SELECTION

Presuppositions in a Selection Program. Personnel selection presupposes differences in worker performance. In preceding discussions the significance of individual differences in job performance has been emphasized. Clearly any program which differentiates among candidates for jobs def-

initely implies that individual differences are important. Specifically, it can be said that whenever a program for the selection of workers exists, it is presumed that individual differences in job performance are recognized, and that the magnitude of the differences in the manner in which various workers perform their jobs is important.

While in most cases differences in the manner in which jobs are performed are important, this is by no means a universal proposition. For example, if a random sample of persons were given some simple task, such as sweeping a sidewalk, the performance of the best worker undoubtedly could be easily distinguished from that of the poorest. If the poorest performance were carefully examined, however, it might be found wholly adequate. There is little need to sweep a sidewalk so clean that not a speck of dirt can be found on it. All that is really necessary is to free it of the larger pieces of litter and most of the dirt, and this might have been accomplished by the least effective worker. It is quite possible that a similar state of affairs exists in many jobs, differences in performance being considered important when in actual fact they really are not. The statement sometimes made that "all that is needed for a job is warm bodies" is one that undoubtedly has some substance.

The widespread notion that differences among workers in terms of job performance are always large is not necessarily true. The statement usually made is that for the majority of jobs the best worker is two to eight times better than the poorest. Such differences certainly would be significant in contributing to production costs when it is remembered that costs for material, overhead, etc., are very nearly the same for all workers on a job. In many instances, however, the performance of the best worker may surpass that of the poorest by as little as 5 or 10 per cent. In some instances, as on a production line or a similarly paced activity, there may be no differences at all in output among the workers. Furthermore, social pressure from superiors or fellow workers may result in little or no variation in rate of production on non-paced jobs. Hence, before a selection program is installed, it is necessary to determine whether differences in job performance exist, and, if so, whether they are large enough to merit establishing a program.

The Two Types of Selection Problems. The problem of selection is twofold: first, differentiating between applicants who possess the requisite qualifications for a job and those who do not possess these qualifications; and, secondly, ranking the applicants from the highest to the lowest in terms of their qualifications. In the first case, the basis of selection is some minimum qualifications that must be achieved or surpassed before the individual is considered satisfactory. In the second case, the basis of selection is the number of workers needed for the job.

Thus, if 20 workers are needed for a particular job and 50 applicants are tested, the 20 earning the highest test scores are selected.

An example of the first type of selection is the situation which exists in large organizations where there are almost always some openings. Any person who has certain minimum qualifications is hired. Persons who do not have these minima, however, are considered poor risks and are rejected. Some of the so-called "competitive" civil service examinations illustrate the second type of selection. In this case, a certain number of vacancies in a particular job exist. On the basis of the test employed, the applicants for the job are ranked in order of their ability. Beginning with the highest-ranking individual and proceeding down the list, that number of persons needed to fill the vacancies is selected.

While these two problems are by no means completely dissimilar, they nevertheless focus attention on two different aspects of selection. In that problem which is concerned with ascertaining whether the applicants possess the minimum qualifications, the implicit assumption is made that, regardless of how job performance is measured, there is a readily determinable critical point above which the individual contributes significantly and below which he is a loss to the organization. The effectiveness of the selective procedure is judged in terms of the degree to which it correctly forecasts which applicants will perform above the critical level and which will not.

In the second problem of selection, which is concerned with ranking applicants in terms of their ability, no assumption is made concerning a critical point in job performance. The assumption is made, however, that any person hired will be at least minimally satisfactory. The task then is simply to select those individuals whose performance on the job will be the best. With this type of problem it is clear that the effectiveness of the selective procedures is gauged in terms of the accuracy with which it differentiates between individuals with respect to their probable success in performing the job.

In many situations the two problems are dealt with simultaneously. Applicants are ranked in order of their potentiality to do the job, but only those are considered for employment who satisfy certain minimum qualifications. Thus, if there are five vacancies and only three out of ten applicants satisfy the minimum qualifications, those three will be hired and two positions will remain vacant. On the other hand, if seven persons qualify, the highest ranking five will be hired and the other two, who also meet the minimum qualifications, will be rejected.

Selection and Individual Growth in an Organization. In considering selective procedures it is necessary to keep in mind whether the appraisal made at the time of hiring is to be restricted to the applicant's capacity to do the immediate job for which he is being considered, or whether

future promise for higher jobs is also to be evaluated. Of the two, measurement of potential growth and development presents the greater difficulties. The development of adequate selective procedures in this field is as important as the development of procedures for the entry job. This is especially true since the general practice in selecting supervisors, minor executives, and administrators is to choose them from persons already employed by the organization in lower positions.

It is obvious that, if in the original selection of workers no attention is paid to the evaluation of potentiality for higher jobs, there is little assurance that sufficient numbers of qualified individuals will be available for such jobs. Where the jobs are under civil service appointments this very problem has been encountered. Owing to a popular sentiment against including in civil service examinations any tests not directly related to the specific jobs for which workers are needed, tests of potentiality for higher order work have been omitted. In many instances, therefore, it has been difficult to get sufficient numbers of employees with the potential ability to develop the qualifications required for the higher supervisory and executive positions.

Promotion and Upgrading as Special Cases of Selection. The problems in promotion and upgrading are primarily the same as those in selection for entry jobs. Jobs are often organized in a hierarchical arrangement, and those lower in the scale are considered necessary stepping stones to the attainment of those higher in the scale. In promotion, one or more vacancies occur at a higher level and the candidates are restricted to those individuals who currently are holding positions at a lower level in the same organization. The task is to rank workers in lower positions according to their qualifications with the objective of filling the vacancies with those of highest ranks. In an office, for example, when the chief clerk's position becomes vacant the senior clerks would be appraised in terms of potentiality to fill the higher position. The most outstanding senior clerk would be promoted.

In upgrading, the situation ordinarily is similar to that existing in selection where there are vacancies and the task is to ascertain whether the individual meets the minimum qualifications. In many skilled trades, the apprentice's progress is appraised from time to time and whenever he meets minimum qualifications with respect to skill and knowledge he is upgraded. This situation may also exist in office work where a worker is upgraded whenever he meets minimum qualifications.

It is apparent from this discussion that promotion and upgrading are simply special cases of the more general problem of selection. Principles and procedures that are found to be useful in initial selection may also be found useful in promotion and upgrading.

Promotion may not be based on potentiality to do the higher job, and

therefore the appraisal of the potential ability of the candidates is not an important problem. Seniority in time on one job is frequently the sole condition for advancement to another job. There are real justifications for the use of seniority as a basis of promotion. It provides a means for more effective utilization of those workers having the greatest experience, it rewards faithful past service, and it is a device for motivating workers to continued efforts. Seniority appears warranted when the knowledge and skill acquired from experience in the lower job are of primary importance for performance in the higher one, when qualifications for the higher position are relatively low, or when the candidates for promotion are about equal in ability to perform the duties of the higher job. When none of these conditions is met, then seniority does not appear to be a wholly satisfactory basis for promotion.

Sometimes potentiality for promotion is determined solely by the proficiency shown on the lower job. Promoting workers on this basis is helpful in motivating them to better performance. This condition is not a sound basis for promotion, however, unless high proficiency on the lower job means that the worker possesses those skills and knowledges required for performance on the higher job.

With the exception of those situations where it has been decided that the function of promotion is to increase motivation and provide rewards, promotion should be based solely upon potentiality to perform the higher work. This potentiality may be best indicated by the amount of experience or proficiency gained on the lower job, by tests, by interviews, or by some other means. But regardless of what procedure is utilized, it can only be justified if the appraisals of potentiality it yields can be demonstrated to be related to performance on the higher job.

Transfer and Reclassification as Special Cases of Selection. In promotion the concern is with the vertical movement of employees. Sometimes it is necessary to deal with problems of horizontal movement. Here employees move between jobs that are approximately equal in terms of such factors as pay, skill, responsibility, and social status. Transfer and reclassification are likely to occur when the need for a particular job ceases. Instead of releasing the workers who held that job, steps are taken to place them elsewhere in the organization. Occasionally personnel are needed for vacancies suddenly occurring in an organization, or for new positions that have been created. For one reason or another it is not desirable or possible to go outside of the organization for new workers so the positions are filled by transfer and reclassification of other workers.

Transfer and reclassification problems are sometimes like problems of selection and sometimes like problems of classification. In situations comparable to those of selection, the workers in jobs at the same level as that of the job in which vacancies exist constitute the available pool of

candidates. Those workers in this pool who are best qualified or who meet the minimal qualifications for the vacant positions are selected and transferred out of their present positions. However, when a particular job is abolished and the personnel are not to be separated from the organization but are to be reclassified and transferred to some other job, the problem is not one of selection. Under such circumstances none of the individuals can be rejected; each must be placed in some job. This situation is like those found in classification. Problems of classification will be dealt with in detail later.

SELECTION AND PREDICTION

Selection as a Problem of Prediction. Selection may involve the ordering of candidates for a job in terms of their capacities to perform the job, or the appraisal of candidates in terms of whether they possess or do not possess the minimum qualifications for the job. In either case, however, the aim is to differentiate the poorer from the better candidates in order to forecast their future success.

The task of selection is one of predicting how well the individual will do on a particular job before he has even worked on it. On the basis of information obtained about the individual by one means or another, such as from his previous occupational history, an interview, or tests, the attempt is made to forecast his behavior in the future, that is, his performance on the job in question once he is placed on it. The best way of finding out what he will do on the job, of course, is to hire him and directly observe his performance. This ordinarily is too expensive to be feasible and, in many cases, is administratively impossible. It is therefore necessary to use some "short-cut" methods as aids in predicting later behavior.

Prediction versus "Description of Qualifications." It is sometimes wrongly believed that the purpose of the appraisal in selection is to obtain a description of the applicant's qualities and capacities. Thus, as a result of the evaluation of a candidate for a particular position he may be described as highly motivated, interested, and having an appropriate training and experience. Because this description is favorable and the candidate is a so-called "good" person he is hired. However, management's interest is not in such a description or appraisal for itself alone. The description has value only in so far as it is known that persons so described are more likely to perform better on the job than those who are not so described.

It cannot be said that individuals who are favorably described necessarily are the most likely to succeed. A person writing a letter of recommendation for another for whom he does not have an especially high

opinion may not wish either to recommend the candidate or to make derogatory statements about his abilities. In such a letter he is likely to describe the candidate as well-appearing, honest, and coming from a good family. This description, while favorable, is probably indicative of poorer rather than better performance on the job. It is also possible that, for a particular job, extensive previous job experience and high education level might be characteristic of those persons who, when hired, turn out to be failures.

When emphasis is placed upon obtaining a description of the individual as he is at the present time, the scope of information obtained concerning him is restricted. Other types of data, which may not appear pertinent but actually may be more predictive, will be overlooked. For example, it may happen that workers who have to make one or more transfers on the bus in coming to an establishment are more likely to stay on the job only a short time. With interest centered on obtaining a description of the person's present qualities and capacities, this important fact may be overlooked. If the description of the individual's present qualities and capacities is helpful in prediction, then and only then should there be any interest in obtaining it.

SELECTION AND THE VALIDITY OF APPRAISALS

The Concept of Validity in Selection. The degree to which a device used in the selection of workers predicts their later success on the job is called its *validity*. A device that has been shown to predict occupational success is said to be a valid one; one that does not is said to be invalid. Some adequate measure of validity is absolutely necessary before the value of a selective device is known and before the appraisals it yields have any meaning as predictors of job success. Actually, the significance and value of a selective device are unknown until it has been checked against adequate, independently determined measurements of job success. Different selective devices will vary in terms of their power to predict a given index of job success. Also a particular device will vary in terms of the validity with which it will predict different indices of job success.

Degrees of Validity. In reference to a given measure of proficiency a selective device should not be described as being either valid or invalid. Rather it should be described in terms of the degree of validity it possesses. Validity is a characteristic which varies from high to low, indicating predictive power from great to little. In response to a given job index one selective device may possess high validity, another only moderate validity, and still another may have no validity at all. A correct description of a selective device, then, would indicate the degree to which

it is valid; in other words, the extent to which predictions concerning future job success are improved over chance by using the information that the device yields.

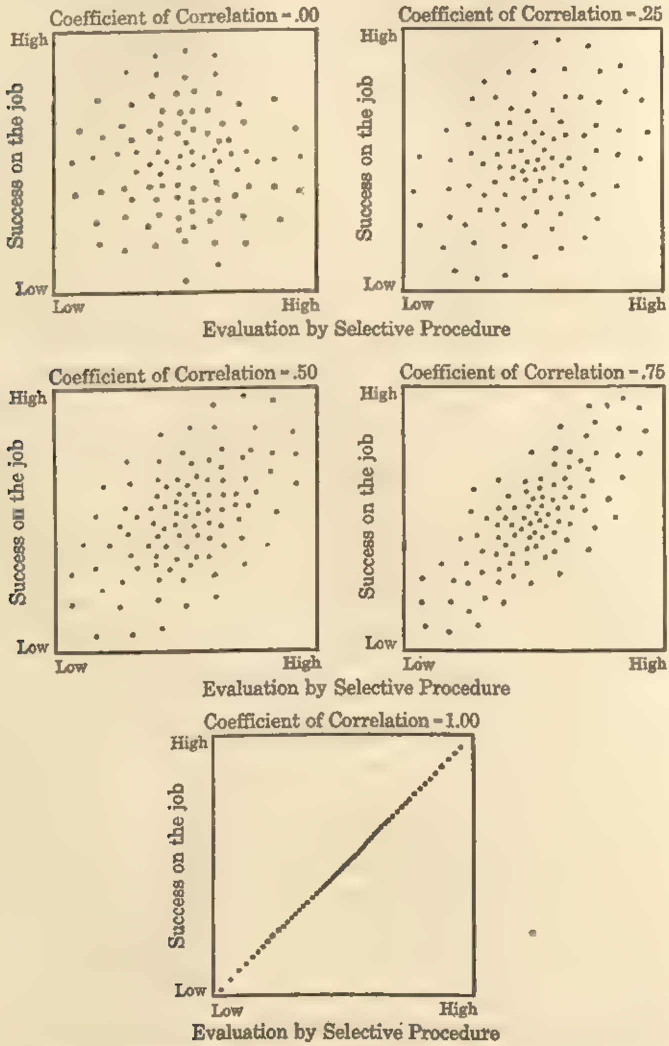


Fig. 5-1. Graphic representations of coefficients of correlation of various sizes.

Representations of various degrees of validity are given in Fig. 5-1. Here are shown plots or scatter diagrams of relationships between appraisals given by selective devices and indices of success on the job, the examples being hypothetical. Each dot stands for a single person. A line drawn horizontally from a dot to the vertical axis will indicate

the value of that individual's standing on the index of job success. A line dropped vertically will intersect the horizontal axis at the value of the appraisal given him by the selective device. When the degree of relationship between appraisals and job success is expressed as a coefficient of correlation it is termed the *validity coefficient*. It will be observed that when the validity is perfect (coefficient of correlation = 1.00), the individual who receives the highest appraisal by the selective device turns out to be the best on the job, the individual who receives the second highest appraisal turns out to be second best on the job, etc. With validity coefficients of lower and lower values this orderly relationship gradually breaks down until with a validity of zero the points scatter at random. Thus the higher the validity the more accurately an individual's performance on the job can be predicted. With a validity of zero predictions are wholly chance, and with a validity of unity predictions are perfect.

Referring Validity to Specific Measures of Job Success. In the discussion of the measurement of job success, it was pointed out that there are many ways of evaluating a worker's proficiency on the job, and that there is seldom perfect agreement between any two or more different measures of performance on the same job. It follows, then, that a given selective device will not necessarily predict equally well each of several measures of job success on a given job. Suppose that success on a job is measured by output and wastage. Scores on a test might be found to be positively correlated with output but not with wastage, or positively correlated with wastage but not with output. The reason for this is that output and wastage are not highly related. This means that a selective device can be a valid predictor of one index of job success and not of another.

In Table 5-1 are shown the validity coefficients of several tests used in the prediction of success in various courses of training for aviation maintenance men.[5] It will be noted that the tests vary considerably in the accuracy with which they predict grades in different phases of training. Thus, the analogies test best predicts grades in the course in electricity, whereas the surface development test best predicts grades in blueprint reading and mechanical drafting. The validity of a selective device, therefore, should be stated in reference to a particular index of job success. The statement frequently made that a particular interview procedure or a particular test is good is meaningless unless it is referred to some definite index of job success. In other words, a selective device is good or bad, *i.e.*, valid or invalid, *not* in general, but only in regard to a given index of job proficiency.

Validity and the Adequacy of the Measure of Job Success. It should be pointed out that in practice a selective device is no better than the

Table 5-1. Validity Coefficients of Several Tests for Grades in Three Basic Courses for Aviation Maintenance Trainees

Test	Course of instruction		
	Shop mathematics	Electricity	Blueprint reading and mechanical drafting
Number series	.37	.15	.24
Analogies	.26	.35	.23
Addition	.39	.23	.08
Mechanical movements	.08	.12	.40
Surface development	.35	.21	.50

measure of job success against which it is validated.[10] If the index used to measure achievement on a job is not itself valid, then the device used to predict it is not useful, even though accurately predicting that index. Suppose, for example, that a test used in selecting sales clerks was validated against grades in a two-day training course. Even if the scores on the test were found to be very highly correlated with such grades, the test would not be useful because the important index of success on this job is dollar volume of sales and not performance in a short and perfunctory training course.

If the index of job success does not adequately cover all phases of job performance, a selective device correlating with it can be considered to predict only those phases of the job that are accurately measured by that index. Furthermore, if a test is known to measure some psychological ability such as arithmetic ability, it should not merely be assumed that it is a valid predictor of a job which requires arithmetic manipulations. Assuming validity is insufficient, the test must be checked against some index of job performance.

The foregoing facts point up the need for making accurate and thorough job analyses in which not only are all of the phases of the job discovered, but also the relative importance of the different phases is determined. Thorough job analyses make possible valid measures of job performance. With adequate measures of job performance available it becomes possible to determine accurately the validity of the predictors being evaluated. Not only can statements be made concerning the degree of validity, but also statements can be made concerning exactly what kinds of behavior are being predicted.

The Relationship between Reliability and Validity. The reliability of the index of job success is also an important factor in the validation of a selective device. Even if the reliability of the selective device itself is very high, the appraisals it yields cannot be expected to show a very high correlation with any measure of job success if that measure is highly unreliable. For example, suppose that the proficiency of a group of clerical workers were measured by means of ratings made by the office manager and the chief clerk. If the ratings made by these two individuals showed no agreement whatsoever, then it should not be expected that a test of clerical ability would show any correlation with an index of job proficiency formed by the average of these two sets of ratings. Therefore, the extent to which the appraisals yielded by a selective device can correlate with a particular measure of job performance is limited both by their own reliability and by the reliability of the measure of job success.

The limiting effects of reliability upon validity are shown in Table 5-2. In this table it may be seen that if the reliability of a selective device

Table 5-2. The Highest Possible Validity Coefficient a Predictive Device Can Have in Relation to the Reliability of the Selective Device and the Reliability of the Index of Job Success

Reliability of the index of job success	Reliability of the selective device					
	.00	.20	.40	.60	.80	1.00
.00	.00	.00	.00	.00	.00	.00
.20	.00	.20	.28	.35	.40	.45
.40	.00	.28	.40	.49	.57	.63
.60	.00	.35	.49	.60	.69	.77
.80	.00	.40	.57	.69	.80	.89
1.00	.00	.45	.63	.77	.89	1.00

were .80 and the reliability of the index of job success were .80, the highest possible validity of the selective device would be .80. If the reliability of the index of job success were .20, then the highest possible validity for the selective device would only be .40. Hence, if the reliability of either the selective device or the index of job success is very low, it is impossible to have a predictive power high enough for any practical use. For example, if the reliability of an interview were zero then it could have no validity at all as an instrument for the selection of

workers. Similarly, if the reliability of an index of job success were zero, it could not be predicted by any kind of selective device.

It should be emphasized that Table 5-2 does not show how high the validity coefficient will be when the reliability coefficients of both measures are known. Rather the table shows the limits above which the validity cannot go. Even if both measures were perfectly reliable, the validity might still be zero.

EVALUATION OF THE EFFECTIVENESS OF SELECTIVE DEVICES

Practical Limits of Predicting Job Success. Individuals who have had little or no experience with statistical studies of the effectiveness of selective devices are likely to presume that the degree of validity of such devices is exceedingly high. Actually, even under the very best of circumstances, the prediction of job success is far from perfect. Generally speaking, the degree of validity is best described as moderately good.[8]

To use selective devices most effectively it is necessary to understand the complexity of the predictive situation. It must be remembered that the behavior of human beings is far more complexly determined than that of any machine. A multiplicity of causes lead to success and failure on the job. They are not solely the function of the abilities a worker possesses, but are directly conditioned by the individual's personality characteristics and motivations. One individual may display high performance on a job due to his loyalty to the organization; another as a result of a strong desire for advancement. One may exhibit poor performance because of sheer inability; another as a result of preoccupation with personal affairs. The work and home environments in which the individual finds himself contain many factors determining his job performance. With an effective leader as a supervisor his performance is likely to be superior; but, despite a good supervisor, it may very well suffer if his wife threatens divorce. These more intangible social and personality factors are difficult to set down in terms of numerical values. Despite the enormity of the problem, however, great strides are being made in the quantifying of these variables. It is expected that the near future will see the discovery of more effective measuring devices in this area.

Recognition of the complexity of behavior has focused attention on the need for getting a more complete coverage of the job performances to be predicted. This has led to the use of several selective devices, which are pooled in such a way as to get optimum prediction. The reader will remember that many of the validity coefficients thus far presented for predictive devices have been only moderate in size. The best single devices have validity coefficients varying from .25 to .50. By using sev-

eral measures optimally weighted in a battery, coefficients of validity have been obtained of the order of .60 to .75. Of course, the development of a selective program entails the construction and validation of several instruments, and requires considerable effort and expense. Unfortunately, few industrial organizations are convinced that this evaluation of the complex human factors in production merits the necessary outlay.

The effectiveness of prediction is conditioned upon still other factors, some of which play an important role in the practical situation. It is obvious that the availability of labor is important. Consider two contrasting situations. One organization must hire 50 workers for a given job, and from a very intense recruiting drive can attract only 55 persons. In an identical situation a second organization attracts 300 applicants. Suppose the job in both organizations is one which must be filled because of some time-limited contract where a deadline in production has to be met. The first organization would be forced to hire many poorly trained workers and to institute a training program to bring them up to a satisfactory level of performance. All of the workers hired by the second organization could be drawn from the high levels of ability if some selective procedure were used for differentiating the applicants in terms of their ability to do the work required. The predictive effectiveness of a selective device is therefore determined in part by the amount of available manpower.

In the following section consideration will be given to several procedures available for measuring the success of a selective program. Factors operating in the practical situation, such as the one just mentioned, will be discussed. Procedures will be outlined by which these practical factors can be incorporated into the selective program in order to increase the predictive effectiveness.

Ways of Describing the Effectiveness of Selective Devices. There are several procedures that can be used for determining the effectiveness of a selective device. When success on the job is indicated in the form of a continuous series of values, such as production, the usual way of determining the validity is the correlational method. As indicated earlier, the validity coefficient measures the closeness of relationship between the index of job success and appraisals given by the selective device. Figure 5-1, which shows graphic representations of various degrees of correlation, permits some interpretation of the effectiveness of prediction. Further use of the correlational method in evaluating selective devices will be described in later topics.

When the variables are not represented in a continuous series other procedures can be used to determine validity. It is not always possible to obtain measures of job success on a continuum in which levels of proficiency are represented as numbers varying from low to high. In lieu

of such a refined scale it may be possible to classify workers into several broad groups such as excellent, good, poor; or simply successful or unsuccessful. A measurement of validity can then be made by calculating the average of the appraisals for each of the groups. For example, if the selective device is a test, then the average score for each group can be determined. If the test is valid, the average score for superior workers will be significantly higher than that of inferior workers. Another technique involves computing the per cent of successful workers earning each score. With a valid test there would be higher proportions of successful workers earning high scores and lower proportions of successful workers earning low scores.

The Factors of Selection Ratio, Job Difficulty, and Performance Standards. The question arises concerning the size of the validity coefficient that can be accepted as indicating satisfactory validity. Because there are several factors determining how effective the prediction represented by a validity coefficient will be, no particular value, or even a range of values, can be categorically given as being minimal values for an acceptable validity coefficient.

If the problem is one of selecting workers on the basis of their aptitude or proficiency, then the size of the validity coefficient which will indicate that the selective device is serviceable will depend on the selection ratio.[13] The selection ratio is the ratio of the number of applicants to be selected to the total number of applicants available. When the ratio is small, that is, when only a small proportion of the applicants needs to be hired, then all workers who are hired can be taken from among those of very high promise. There will be no need of selecting any workers from among those of average or low promise. When the selection ratio is high, that is, when most of the applicants must be hired, very little selection can actually be accomplished. Even applicants of mediocre ability must be hired.

In addition to the selection ratio, consideration must be given to the level of performance that is considered adequate according to the difficulty of the job.[6] One way of stating the difficulty of a job is in terms of the proportion of workers who can perform the job in an adequate way, when these workers come from a random group of applicants and are not selected on any systematic basis. A job on which only 20 per cent of such persons can perform in an adequate way is a difficult job. A job on which 80 per cent can perform adequately is an easy one.

To make statements such as the foregoing, that only a certain proportion of workers can perform a job in a satisfactory manner, implies that some standard of performance exists. Those workers whose performance exceeds this standard are said to be satisfactory, and those whose per-

formance falls below it are considered unsatisfactory. A good selective device, then, would have the following characteristics: it would select the highest possible number of successful workers and reject the highest possible number of unsuccessful workers, and it would select the smallest

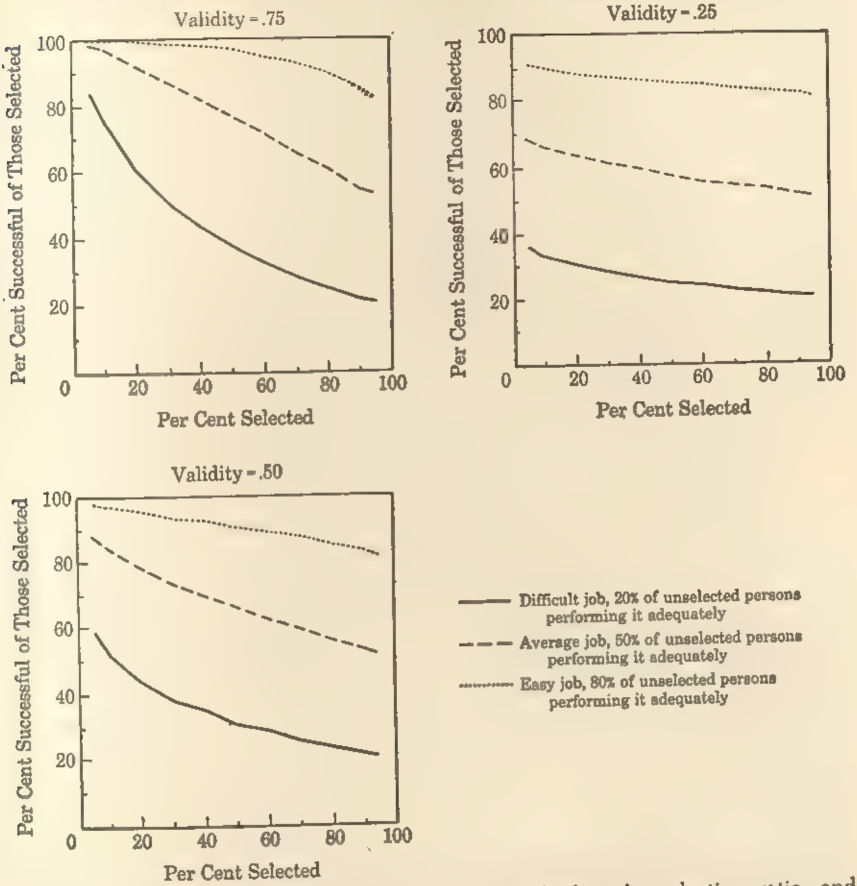


FIG. 5-2. The relationships among validity, job difficulty, the selection ratio, and yield in terms of per cent of selected workers who are successful.

number of unsuccessful workers and reject the smallest number of successful workers.

Evaluation of Selective Devices in Terms of the Per Cent of Selected Workers Who Are Successful. The per cent of selected workers who turn out to be successful on the job is conditioned by the relationships among the factors of job difficulty, the selection ratio, and validity. These relationships are illustrated in Fig. 5-2. From the trends shown in this figure the following principles can be evolved:

The higher the validity of the selective device the larger will be the proportion of selected persons who turn out to be satisfactory (*e.g.*, with a selection ratio and job difficulty of 50 per cent, a validity of .75 will yield 77 per cent satisfactory workers. Under these same conditions a validity of .25 will yield only 58 per cent satisfactory workers).

The smaller the proportion of candidates selected the larger will be the proportion of selected persons who turn out to be satisfactory (*e.g.*, with a validity of .50 and a job difficulty of 50 per cent, a selection ratio of 20 per cent will yield 78 per cent satisfactory workers, whereas a selection ratio of 80 per cent will yield only 57 per cent satisfactory workers).

The easier the job the larger will be the proportion of selected persons who turn out to be satisfactory (*e.g.*, with a validity of .50 and a selection ratio of 50 per cent, a job with a difficulty level of 80 per cent will yield 91 per cent satisfactory workers, but one with a difficulty of 20 per cent will yield only 31 per cent satisfactory workers).

A selective device of lower validity when coupled with a lower selection ratio may be as effective as, or more effective than, a device of higher validity which is coupled with a higher selection ratio (*e.g.*, for a job of 50 per cent difficulty, with a validity of .25 and a selection ratio of 20 per cent the yield of successful workers will be 64 per cent. In a job of equal difficulty, but with a validity of .75 and a selection ratio of 80 per cent, the yield of successful workers will be 61 per cent).

A selective device of lower validity when applied to an easier job may be as effective as, or more effective than, one of higher validity which is applied to a harder job (*e.g.*, with a selection ratio of 50 per cent, a validity of .25 and a job difficulty of 80 per cent the yield of successful workers will be 86 per cent. With a selection ratio of 50 per cent, a validity of .75 and a job difficulty of 20 per cent the yield of successful workers will only be 37 per cent).

Evaluation of a Selective Device in Terms of the Per Cent of Applicants Correctly Placed. Evaluating a selective device in terms of the proportion of successful workers selected is not always the most effective procedure. Frequently the interest is not only in obtaining the highest proportion of satisfactory workers, but also in losing through rejection the smallest proportion of satisfactory workers. This is particularly true in a tight labor market where the loss of successful individuals through their failure to pass a selective program might be as undesirable as the hiring of unsatisfactory persons. The problem can be stated as one of cost and utility: specifically, how many satisfactory workers will be rejected in the attempt to select a group of workers almost all of whom are satisfactory.[15] As the selection ratio is set higher and higher not only will an increasing proportion of those selected be satisfactory, but also an increasing proportion of those rejected will be satisfactory. In other words, by selecting only the few very best candidates there are great gains in terms of obtaining a high proportion of successful workers, but there also are high costs in terms of losing a high proportion of potentially successful workers. Another way of evaluating the effectiveness of a selective program, therefore, is in terms of the number of individuals

it correctly places; that is, the percentage of successful workers it selects plus the percentage of unsuccessful workers it rejects.

The relationships among the characteristics of job difficulty, the selection ratio, validity, and the yield in terms of the per cent of applicants

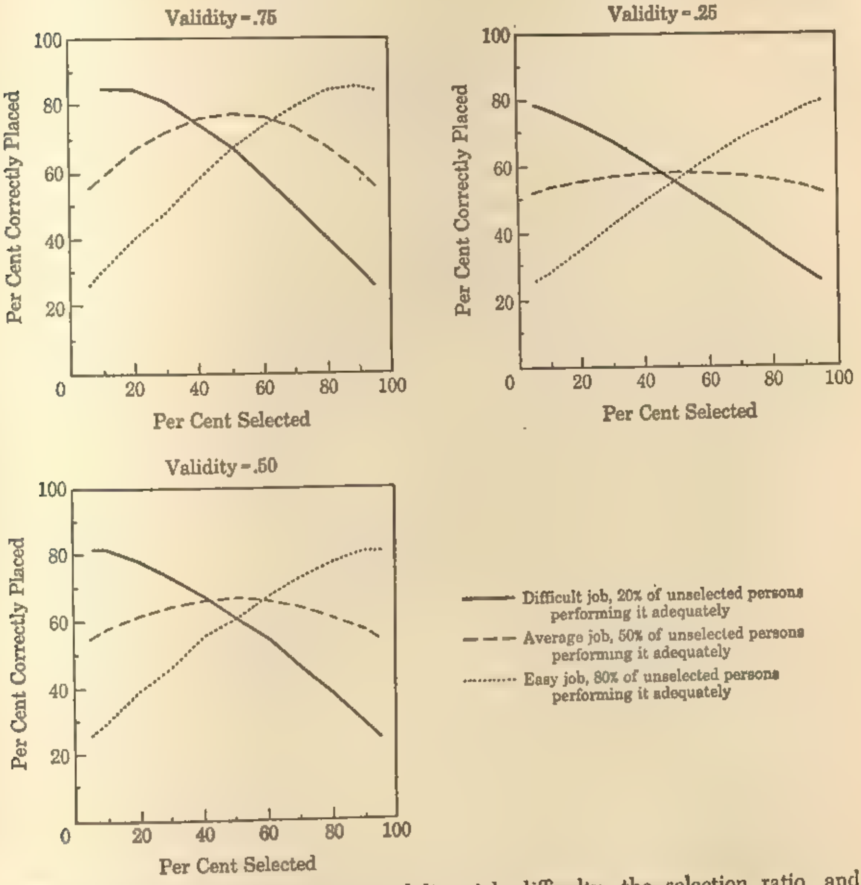


FIG. 5-3. The relationships among validity, job difficulty, the selection ratio, and yield in terms of per cent of persons who are correctly placed.

who are correctly placed are shown in Fig. 5-3. From the trends shown in this figure the following principles can be evolved:

The higher the validity the larger will be the proportion of persons correctly placed (*e.g.*, with a selection ratio of 50 per cent and a job difficulty of 50 per cent, a validity of .75 will result in 77 per cent of workers correctly placed, whereas a validity of .25 will result in only 58 per cent of workers correctly placed).

In general, the more the selection ratio is similar to the per cent of unselected persons who can adequately perform the job, the larger will be the proportion of workers correctly placed (*e.g.*, with a validity of .50 and a job of 20 per cent difficulty,

a selection ratio of 20 per cent will result in 78 per cent of workers correctly placed, a selection ratio of 50 per cent will result in 61 per cent of workers correctly placed, and a selection ratio of 80 per cent will result in only 38 per cent of workers correctly placed).

Evaluation of Selective Devices in Terms of Per Cent Improvement in Proficiency. In many, if not most, job situations it will be impossible or unnecessary to set precise standards of job proficiency. The interest will not be in the increase in the proportion of persons correctly selected or placed but in the per cent of improvement to be expected in some index of job proficiency, such as production, when workers are systematically selected. It will be important, then, to know how output would increase if workers were selected by a device of a given validity using a particular selection ratio.

In order to make estimations of improvement in job proficiency it is necessary to take into account the relative variability of workers in job proficiency.[9] What is meant by relative variability? Suppose in a group of workers the production of the best worker is 150 and that of the poorest is 50. The ratio of the best to the poorest is then 3 to 1 and the relative variability is large. If, however, the production of the best is 101 and the poorest 99, then the ratio of the best to the poorest is 1.02 to 1.00 and relative variability is small. Clearly, in the second case the workers are quite homogeneous with respect to the proficiency they manifest on the job; the range of individual differences among them is very small.*

Figure 5-4 shows the relationships among validity, the selection ratio, variability in job proficiency, and per cent improvement in proficiency. While this chart may seem complex at first glance, it actually is not difficult to read. Suppose it was desired to estimate how much increase in production would result if 10 per cent of candidates were selected with an instrument having a validity of .60, with the ratio of the best to the poorest worker in job proficiency at the present being 2 to 1. Locating the 10 per cent point on the scale of per cent selected (on the portion of the chart to the left), the line is followed up to the point where it intersects the curve for a validity of .60. Then following straight across to the right to the vertical scale for ratio of best to poorest worker of 2 to 1, the per cent increase in proficiency can be ascertained. The proficiency of the selected cases would be approximately 12 per cent better than that of present unselected workers.

* For purposes of simplicity variability is expressed here as the ratio of the best to the poorest worker. A much better index of variability would be the standard deviation. In the present discussion the proficiency of the best worker is taken as falling at three standard deviations above the mean and that of the poorest as falling at three standard deviations below the mean.

The chart can similarly be read for the expected improvements associated with different degrees of variability ranging from quite homogeneous groups (ratio of best to poorest of 1.1 to 1) to extremely heterogene-

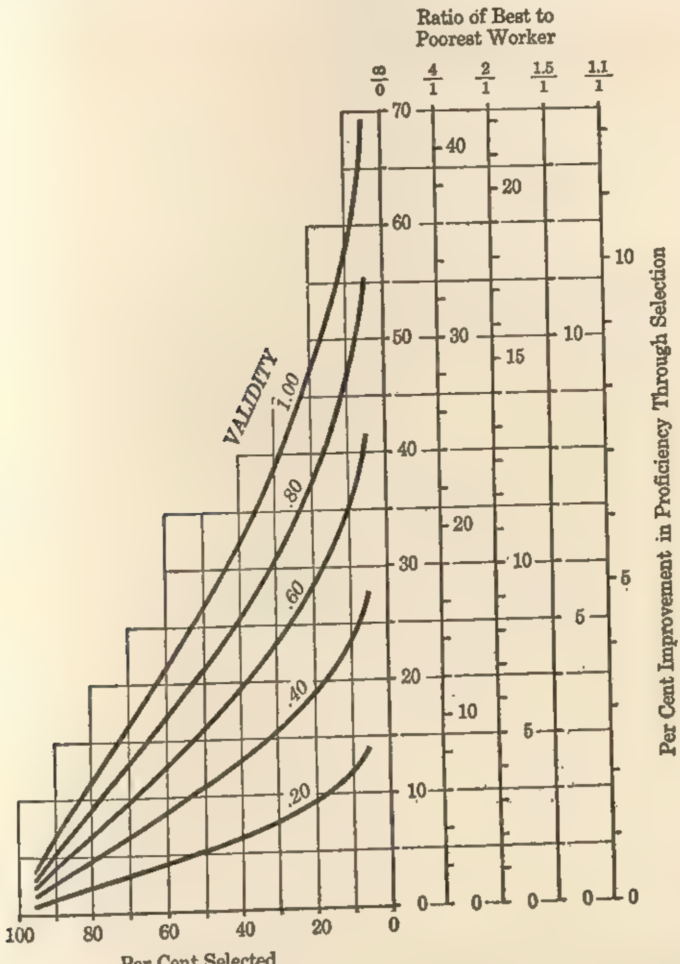


FIG. 5-4. The relationship between validity, the selection ratio, variability in job proficiency (ratio of the best to the poorest worker), and per cent improvement in proficiency.

ous groups (ratio of best to poorest of ∞ to 1), that is, infinitely large to one. The latter would be used when the proficiency of the best worker is 200 and the proficiency of the poorest is one.

From the trends shown in Fig. 5-4 the following principles can be evolved:

The higher the validity the greater will be the average proficiency of those workers who are selected (*e.g.*, with a selection ratio of 50 per cent and a ratio of best to poorest worker of 4 to 1, a validity of .80 will yield an increase in proficiency of 13 per cent, while a validity of .20 will yield an increase of only 3 per cent).

The lower the selection ratio the higher will be the average proficiency of those workers who are selected (*e.g.*, with a validity of .60 and a ratio of best to poorest workers of 4 to 1, the average production of the best 20 per cent of workers as indicated by the selection device will be 17 per cent higher than unselected workers, while the production of the best 80 per cent will be only about 4 per cent higher).

The more homogeneous the workers are with respect to job proficiency the smaller will be the improvement resulting from a systematic selection of workers (*e.g.*, with a validity of .60 and a selection ratio of 50 per cent, when the ratio of the best to the poorest worker is only 1.1 to 1 the increase in proficiency above that of unselected workers will be only 3 per cent. When the ratio of the best to the poorest worker is 4 to 1 the increase in proficiency will be 10 per cent).

CLASSIFICATION AND DIFFERENTIAL PLACEMENT

The Increasing Importance of Personnel Classification. Before World War II the problem of the placement of workers on jobs was thought of as being simply that of differentiating between those workers more likely to be successful and those less likely to be successful. The former were accepted and hired and the latter rejected. The rejected were left to find work elsewhere or to solve their economic problems in other less socially acceptable ways. With the advent of the war, the demands for manpower in both industry and the armed services greatly increased, and the individual became a much more important factor. No longer could he be simply rejected, but rather any capacities he had were necessarily put to use. The task became one of finding that job for him where he could make some kind of positive contribution. As long as the demand for manpower remains at a high level the task will continue to be one of placing every individual in some kind of constructive activity. The problem, then, is one of classification of personnel, rather than one of selection, with its consequent rejection and waste of manpower.

It is expected that problems of classification will become more important in the future. Present indications point to an increasingly tighter labor market. If standards of living are to continue to increase the nation must achieve greater and greater productive power. To accomplish this the potentialities of workers must be more fully realized. In part this can be attained by placing people on jobs where their contribution to the general welfare will be greatest. In addition, marginal workers and those previously considered as unemployable must be placed in useful activities. It appears that the future will bring an increased insistence that all individuals be dealt with in a positive fashion so that each is placed in that type of work where he can contribute his share of con-

structive advancement to himself, to his employer, and to the community. It may, therefore, be expected that there will be a greater concern for an individual's relative fitness for different types of jobs than for the absolute amount of his abilities.[16] This problem is one of classification.

Placing the Worker in the Job for Which He Is Best Qualified. As the classification problem has been discussed in the preceding topic, the reader is likely to have the impression that the aim of classification is simple. It might be thought that the objective is the placement of each individual in that job for which he is best fitted. The task would be to discover for each person his greatest capacities and to place him in that job where those abilities are the important ones. Surprisingly, this is not the objective.

Placing every individual in the job for which he has the highest capacities may result in jobs being filled by unqualified workers. The situation can be illustrated by the hypothetical case presented in Table 5-3. Sup-

Table 5-3. Illustrations of Types of Assignments Made in Classification in Relation to Effectiveness of Placement

	Job A	Job B	Job C	Number of jobs adequately filled	Number of workers placed in accordance with their highest talent
Requirements of 3 jobs expressed in terms of % of persons who cannot adequately perform them	90	50	10		
Abilities of 3 workers expressed in terms of % of persons with less ability:					
Worker I	90	50	99		
Worker II	80	70	5		
Worker III	20	25	15		
Types of assignments:					
Placing workers in accordance with their greatest talents	Worker II	Worker III	Worker I	1	3
Placing workers so all jobs are held by persons with adequate talent	Worker I	Worker II	Worker III	3	0

pose there are three workers who differ in their abilities and one of the three is to be placed in each of three jobs, the jobs having different requirements. For purposes of illustration the abilities of each individual are expressed in terms of the per cent of persons who possess less ability. Thus worker II is fairly high in the abilities required for job A, surpassing 80 per cent of people; almost as high in the abilities required for job B, surpassing 70 per cent of people; and quite deficient in the abilities required for job C, surpassing only 5 per cent of people. The job requirements are expressed in a similar way. Job A is a difficult job since 90 per cent of workers cannot perform it in an adequate fashion, job B is one of average difficulty since 50 per cent of workers cannot perform it adequately, and job C is a very easy job since only 10 per cent of workers cannot perform it.

Now from Table 5-3, it can be seen that if each worker is placed in that job for which he has the greatest talent, worker I will be placed in job C, worker II in job A, and worker III in job B. If this type of assignment of the individuals is made, then each person is assigned to that type of work for which he has the greatest ability. It will be noted, however, that only one of the three jobs is filled by a qualified person. Worker I is much more than minimally qualified for job C, but neither worker II nor III can adequately perform the job to which he is assigned. From this analysis it is apparent that the classification problem cannot be stated merely as the placing of each worker in that job for which he is best fitted.

A fundamental difficulty with this statement of the objective of personnel classification is that it presupposes that for each worker there is one job which he can perform best. The situation as presented in Table 5-3 is far too simple. The presupposition is that a worker's abilities and traits are so organized that there is only one job which he can do best. This is the old and psychologically invalid notion of the "perfect niche." The facts of the matter are that a person may be able to do two or more things equally well. A man who is a top-notch drill-press operator undoubtedly can operate a punch press almost equally well. Similarly, a college student who does well in a major in history probably can achieve the same grades in sociology or political science. To be sure, the drill-press operator might be a poor assembler, and the history major might fail in engineering. But for both persons success is not necessarily restricted to one type of activity. For each the breadth of abilities is sufficiently great, and the scope of activities sufficiently broad, to permit of about equal success in several different fields of endeavor.

Placing the Worker in That Job Which Will Effect the Greatest Group Achievement. Referring again to Table 5-3, suppose that each of the three workers is assigned to a job for which he is qualified without giving

special attention to the worker's highest abilities. Let worker I be assigned to job A, worker II to job B, and worker III to job C. It will be seen that every job is filled by an individual who can perform it adequately. With each job being performed adequately it could be assumed that the greatest effectiveness of the group is being realized. It should be observed, however, that no person is doing that kind of work for which he is best qualified. Workers I and II are assigned to jobs for which their talents are secondary, and worker III is assigned to a job for which his qualifications are lowest.

The statement of the objective of classification that now emerges is that individuals should be placed in jobs so that maximum effectiveness of operation of the total organization results. In order to achieve this end it will be necessary in many cases, if not most, to ignore a given person's greatest individual talent. The unit of consideration is not the individual but the total working group.

There are some very important difficulties with this second concept of personnel classification. First of all the social values in a democracy do not permit individuality to be completely submerged. While it is true that certain types of limitations are placed upon individual action, such limitations are introduced in order that the individualities of the largest numbers of persons can be realized. In a democracy the objectives of group endeavor are not final goals in and of themselves. Rather they are considered as means for developing and strengthening the individual goals of the group members.

A further difficulty with this concept of classification stems from the implicit idea that the effectiveness of an organization can be readily determined and measured. One widespread notion is that effectiveness can be judged solely in terms of net productivity. There is no question but that high production is an important objective, but certainly it is not the only one. The occupational placement of people so that they can realize their greatest satisfaction also is important. The criteria on the basis of which the effectiveness of group operations is to be judged are complex rather than simple. As was pointed out in the chapter on the measurement of job proficiency, an individual worker's success ordinarily cannot be measured in a single dimension. This same argument is applicable to the effectiveness of performance of a group which can be manifest in many ways. The contributions of each member of a group to each of the group's achievements may vary widely. It is not a simple problem to fit the individual into the group in a way that will achieve the greatest effectiveness of the group.

From the foregoing discussions it should be apparent that it is by no means easy to state the problem of personnel classification. The objectives are not at all clear. Classification does have to do with the placement

of individuals in the jobs within an organization and does emphasize the need for achieving the organization's goals. At the same time the individuality of the worker also must be considered. At present there is no clear-cut procedure for resolving the conflicting issues.

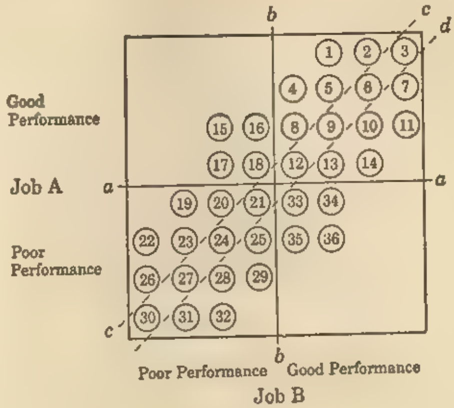
Classification as a Function of the Correlation between Performances on Jobs. In the absence of well-established general principles for the classification of personnel, it will prove profitable to consider specific problems involving the assignment of workers to jobs. Through such problems some of the factors affecting classification and differential placement will be pointed up and possible arguments in support of different solutions may be discovered.

Ease of assignment to jobs is a function of the correlation between performances on the jobs involved. When the correlation between jobs is positive it is more difficult to make assignments than when the correlation is negative. Stated in other words, when abilities and traits required for success on one job are very similar to those required on another, the differential assignment of workers to the two jobs is a difficult task. On the other hand, when the abilities and traits that lead to success in one job are those that lead to failure in the other, the task of differential assignment is relatively easy.

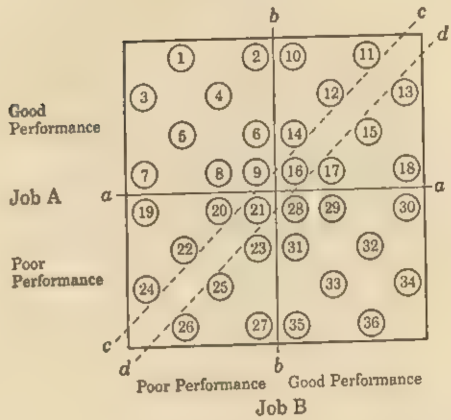
A specific problem in classification involving variation in the correlation between job performances can be illustrated by a hypothetical example requiring the assignment of workers to one or the other of two jobs. With increasing positive correlation between performances on the jobs the classification becomes increasingly more difficult. With increasing negative correlation, classification becomes increasingly more easy. The hypothetical example is illustrated in Fig. 5-5. Suppose that the problem is to assign an equal number of workers from a pool of 36 candidates to each of two jobs, A and B. In each of the examples in the figure job A is represented on the vertical axis and job B on the horizontal axis. Each encircled number represents a worker. A given worker's performance in job A can be determined by moving from his circle on a level line to the vertical axis. Similarly, his performance in job B can be determined by dropping a vertical line from his circle to the horizontal axis.

Example I in the figure represents the situation when performances on the two jobs are positively correlated. In order to get the best 18 men for job A, all those above line *aa*, numbers 1 through 18, would be assigned to it. Similarly, to get the best 18 men for job B, all those to the right of line *bb* would be assigned to it. It will be noted, however, that 14 of the men, numbers 19 through 32, would be discarded. For only 8 men would the assignments to jobs be unique, numbers 15 through

Example I
when performances on the two jobs are positively correlated



Example II
when performances on the two jobs are uncorrelated



Example III
when performances on the two jobs are negatively correlated

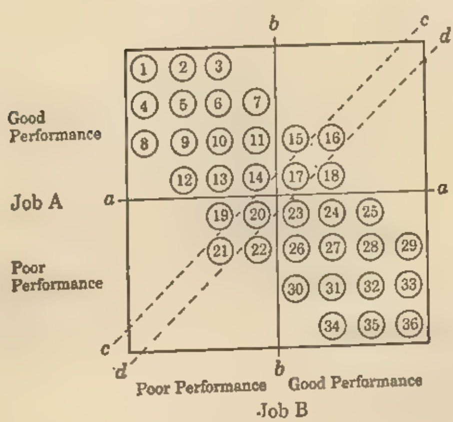


FIG. 5-5. Examples of ways in which 36 persons can be assigned to two jobs.

18 who would be assigned only to job A and numbers 33 through 36 who would be assigned only to job B.

Referring now to Examples II and III in Fig. 5-5, it will be seen that many more unique assignments can be made. In Example II, where performances on the two jobs are uncorrelated, 18 men can be uniquely assigned. Numbers 1 through 9 would be assigned only to job A, and numbers 28 through 36 would be assigned only to job B. When performances on the two jobs are negatively correlated, as in Example III, then 28 unique assignments can be made. Numbers 1 through 14 would be assigned only to job A, and numbers 23 through 36 would be assigned only to job B. In Example III it can be clearly seen that the workers who can be given unique assignments to one or the other of the two jobs are those whose performance on one job is high and whose performance on the other is low. This state of affairs is best realized when the correlation between performances on the two jobs is negative and high.

The problem of placing those workers whose assignments obviously are not unique will now be considered. In Example III of Fig. 5-5 the assignments are not clear for eight workers, numbers 15 through 22. Numbers 15 and 19 perform a little better on job A than on job B, while for numbers 18 and 22 the reverse is true. Thus it would seem advantageous to assign the first two individuals to job A and the last two to job B. Numbers 16, 17, 20, and 21 perform about equally well on both jobs, and hence these workers could be assigned to the jobs at random, *e.g.*, 16 and 21 to job A and 17 and 20 to job B. If this procedure is applied to each of the three examples it will be seen that all workers falling above line *cc* perform better on job A than on job B and hence should be assigned to job A. Similarly all workers falling below line *dd* perform better on job B than on job A and hence should be assigned to job B. Workers falling between lines *cc* and *dd* are equally good on both jobs and should be assigned randomly. In the above assignments account is taken of each individual's highest talents in so far as it is possible.

To generalize tentatively from a situation where there are only two jobs to one where there are many, it can be concluded that unique classification or unique differential placement is accomplished when each individual shows outstanding talent for one job and little or no talent for the remaining jobs, and when performances on the various jobs are negatively correlated. Two circumstances make it very unlikely that either of these conditions will occur. While there are some individuals who show single outstanding talents, and mediocrity or deficiency in all other respects, they are extremely few in number. Most persons possess a wide range of abilities and traits, which, when distributed with respect to

amount, tend to form a continuous series extending from outstanding to very poor.[7] Furthermore, seldom are tasks found which are negatively correlated, that is, where individuals who perform well on one task perform poorly on another. From hundreds of empirical studies the usual finding is that performances on different tasks either are uncorrelated or are correlated positively.

In the classification scheme presented in Fig. 5-5 each of the two jobs received about the same amount of talent in its assignments from the 36 candidates. The unique assignments to the two groups were equal in number and represented comparable levels of talent. In the assignments which were not unique each job received the same number of workers and comparable amounts of talent. In each example of the figure some assignments had to be made at random. Random assignment, however, would result in each job getting about the same amount of talent. Neither job received the best workers and a compromise was reached in which there was an equitable division of talent between the two jobs.

It may not be desirable, however, to have an equitable distribution of talent. Suppose job A was that of assembler and job B that of inspector. It might be desirable or even necessary in a particular situation to have the best possible inspectors even at the price of having poorer assemblers. It is obvious that to select the best possible workers for the one job might result in a sacrifice in talent for the other. In this case, all workers who fall to the right of line *bb* in Fig. 5-5 would be placed on job B and those to the left on job A. When performances on the two jobs are positively correlated, as in Example I, the workers who are to the left of line *bb*, and therefore assigned by default to job A, would be poorer in performance than if a random assignment of the 36 candidates to the two jobs were made. When the two performances are uncorrelated, as in Example II, there would be no loss in talent in job A from what would be expected from a random assignment of all workers to the two jobs. Of course, there would be no gain in talent either. Greater talent than that ensuing from random assignment of the candidates would result for job A when the two performances are negatively correlated as in Example III. As previously indicated, however, negative correlations between human abilities are extremely rare. Procedures of this type for classification of personnel have long been known but seldom employed in any systematic fashion.[14]

The foregoing discussion will serve as a basis for making some specific statements about the aims of classification and differential placement. The end result desired for any given job or for all jobs might be any one of the following:

1. The highest possible average performance
2. Some particular average performance

3. The highest possible proportion of persons who meet some given standard of performance
4. Some particular proportion of persons who meet some given standard of performance

For example, suppose that men were to be placed in the following jobs: assembler, inspector, packer, and maintenance. It might be desired simply to place the men so that the highest possible average performance results for each job. However, some jobs may be more important than others or require higher standards of performance.[1] Therefore, some more complex goal might be desired such as the placement of men on the job of assembler so that on the average they produce 10 units per hour; the placement of men on the job of inspector so that each one achieves an accuracy of at least 90 per cent; the placement of men on the job of packer so that the highest possible proportion pack at the rate of 10 units per hour or better; and the placement of men in the maintenance job so that on the average their performance is as high as possible. The extent to which these results can be achieved will depend upon a number of factors, some of which will be described in the following paragraphs.[15] In some situations it may not be possible to achieve the desired results because of the caliber of the personnel available or the particular nature of the jobs involved.

Classification as a Function of the Nature of the Relationship among Jobs. In the preceding topic classification was considered in reference to the correlations existing among job performances. Another important factor affecting classification is the nature or type of the working relationships that exist among the jobs to be filled.[4] Jobs fall into three different types with respect to the nature of their interrelationships, namely, independent jobs, jobs that are performed successively in a sequence, and jobs that are coordinate. For each relationship the classification problem is different.

The earlier discussions of classification are primarily applicable to independent jobs, that is, jobs that contribute independently, or at least relatively independently, to the effectiveness of the total group or organization. A large number of jobs are of this kind. In a department store, for example, the work of a sales clerk in one department generally will have relatively little bearing upon that of a sales clerk in another department. Here the classification problem is seen in its simplest form. The task is much like that described in the examples of Fig. 5-5. In so far as possible each person is placed in that job for which he has the greatest talent and at the same time an attempt is made to achieve an equitable division of talent among the several jobs.

With jobs that succeed each other in a sequence the problem of differential placement is more difficult. In this case the work passes from

one individual to another. The performance of those whose activities occur late in the sequence is directly conditioned by the performance of those who play a part earlier in the sequence. For example, in a bottling operation a bottle will first be handled by the filling-machine operator, then by the capping-machine operator, and finally by the packer. If the filling-machine operator is very slow, then the production of the other two workers is restricted. Under these circumstances the most effective procedure would be to assign workers to the three jobs so that the average speed of production in each is exactly the same. The problem here is one of allocating talent to each job to maintain a flow of production that is smooth, orderly, and at a maximum rate.

Finally, classification must deal with jobs that are coordinate in nature. Here some end product results from the joint effort of the members of a group, each member of which performs a different function. This essentially is the problem of forming the best team with a given number of individuals. A riveter and a bucker-upper furnish a good example. No matter how good the riveter is, if the bucker-upper fails to hold the rivet securely, the work will be poorly done or done slowly. Another example of a team is the crew of an airplane, where the pilot, copilot, flight engineer, and navigator each perform a different function. Their activities are coordinate in the sense that the effectiveness with which their craft is operated is a function of their joint activities. An important characteristic of coordinate jobs is that they involve interpersonal relationships of various kinds. Hence consideration must not only be given to individual talents, but also to interpersonal reactions. Thus workers A and B might be a good team, as also might workers C and D. But worker A teamed with worker D might be a failure, even though workers C and B could work together effectively.

Validity of Predictors in Classification as a Function of the Relationship among Jobs. In classification the effectiveness of predictors is a function of the nature of the relationships among the jobs to which assignments are to be made. Thus far in the discussions of classification the presumption has been made that information is at hand concerning the proficiency of each individual on every job. This situation, however, is a most unlikely one. Seldom will a situation exist wherein a given individual will be completely trained for each of a number of different jobs and will have performed on each to the extent that his proficiency on every one is known. What is necessary, therefore, is to use predictive devices to estimate his job performances. As was pointed out earlier, in selection the most desirable characteristic of predictive devices is high validity, that is, a high correlation between the assessments of worker aptitude and performance on the job. In classification, however, such a generalized statement cannot be made. The effectiveness of predictors

must be separately evaluated for each of the three types of classification problems discussed above, namely, independent jobs, successive jobs, and coordinate jobs.

With independent jobs the task is not the determination of the absolute levels of aptitudes of a given worker for the different jobs. Rather, it is the determination of the differences between his aptitude levels for the different jobs.[15] The problem can be illustrated with a simple situation involving two jobs, A and B. The information needed is not how much talent a worker possesses for job A and for job B (ability A and ability B). Instead the information needed is how much *more* talent he has for the one job than the other (ability A *minus* ability B, or B *minus* A). A good predictive device in this situation, therefore, does not necessarily have high validity for each job. The characteristic it does possess, however, is power to differentiate between abilities in the two jobs. For example, for purposes of this type of classification, a test which has a validity coefficient of .70 for job A and .65 for job B would not be as effective as one which has a validity coefficient of .30 for job A and .05 for job B. In the latter case the difference between the validity coefficients is greater. With independent jobs, then, the general principle is that good predictors should show wide differences in validity for the various jobs involved. Absolute level of validity is less important.

In differential assignment to jobs related in a functional sequence, absolute level of validity is important. Again the problem can be illustrated with two jobs. Here it is desired to place on the two jobs workers whose abilities bear a known relationship to each other. Suppose the flow of work is such that it is necessary to place on the two jobs workers of equal productivity. Then for each worker who has the ability to turn out 10 units of work per hour on job A another worker must be placed on job B who has the ability to turn out 10 units per hour. In general the higher the validities of the predictors the more accurately will the abilities to do the two jobs be estimated, and the more accurate will be the pairing of workers on the two jobs. A further desirable condition is that the degrees of validity for the two jobs be about the same. If the validity for one job greatly differs from the validity for the other, then the matching of workers on the two jobs will be more difficult.*

The problem of deciding what kinds of relationships among predictors will facilitate assignment to coordinate jobs is a difficult one. The problem has not yet been stated with any great degree of clarity. In fact, for coordinate jobs it may be impossible to state the effectiveness of a predictive device in terms as simple as a validity coefficient. Again, for

* The student of psychological measurement will recognize that with variation in the validity coefficients, regression effects will be different. To some extent errors of matching can be minimized by taking these effects into account.

purposes of simplicity, suppose only two jobs are involved. The question to be answered is, what are the characteristics that should be possessed by the worker placed on job A and the characteristics that should be possessed by the worker placed on job B, so that the performance of the team of two men is at a maximum? The probability is that in most cases there will be a number of sets of characteristics that will be equally good. In the case of the riveter and bucker-upper it might be that, in terms of production, equally effective teams could be formed by having two strong men, or having two men high in arm dexterity, or having one high in arm dexterity and the other strong.

When consideration is given to the use of placement devices with respect to the classification of workers on coordinate jobs, many possible combinations of talents emerge. To give some indication of the range of possibilities, a few simple combinations are listed below. For purposes of illustration it is presumed that there are only two jobs, and measurements of abilities are made by tests. Maximum team production might result when:

1. Both members of the team have equal test scores.
2. The test score of the person placed on job A is superior to that of the person placed on job B.
3. The score on test *a* earned by the person placed on job A is equal to the score on test *b* earned by the person placed on job B.
4. The person placed on job A earns a high score on test *a*, and the person placed on job B earns a high score either on test *b* or *c* but not both.
5. The person placed on job A earns a higher score on test *a* than on test *b*, the reverse being true for the person placed on job B.

The reader could think of many similar possibilities. The ultimate problem is to discover which combination of talents, if any, predicts maximum team performance. In general the approach would have to be by trial and error. Many different combinations of characteristics would have to be studied using various devices such as interview ratings, physical measurements, aptitude and achievement test scores, etc. Eventually one or more combinations should emerge which would predict team success.

As organizational activities become more and more complex emphasis on coordinate jobs will increase since it will be impossible for one individual to carry all of the responsibilities and duties connected with a given activity. Thus at higher levels of management, a person in a position of high responsibility is aided by a staff—a group of men functioning as a team. Very little systematic research has been done in the area of the formation of effective teams, and many important problems await investigation. One notable example in team formation was made during World War II in the formation of bomber crews.[2] The hypothesis was made that the most effective team would be formed of persons with

similar interests and personality characteristics and somewhat complementary aircrew aptitudes. A comparison was made between the effectiveness of crews whose members were considered well matched in terms of abilities, age, section of country, religion, etc., and crews whose members were considered poorly matched. Matching was not accomplished in terms of absolute levels in the characteristics but in terms of the relative levels among the members of each crew. Of the well-matched crews 83 per cent were judged to be effective team groups, while only 47 per cent of the poorly matched crews were so judged.

ADMINISTRATIVE PROBLEMS IN SELECTION AND CLASSIFICATION

Effectiveness in Relation to Resources. Relatively few organizations fully exploit their programs of selection and placement. Whereas time, effort, and money are liberally spent in the development of effective methods of work and in the procurement of good equipment, ordinarily few resources are devoted to establishing and maintaining sound personnel procedures. In practice, the talents of men are frequently given less attention than the capacities of machines. Yet machines must be operated by men. It is impossible to estimate whether a given amount of money will bring in greater returns when spent on the physical phases of work than when spent on the human phases. It is certainly clear, however, that well-developed selection and placement programs far more than pay for themselves.

Selection and Classification as Technical Processes. Even when steps are taken to install and operate systematic selection and placement procedures, such procedures often are placed under the direction of individuals who are untrained in the required specialized techniques. Just as the development of methods of work, the effective utilization of industrial equipment, etc., are technical processes and require specialized personnel, so the development and operation of selection and placement procedures require specialized personnel. The effectiveness of such procedures is directly conditioned by the degree to which adequately trained persons are utilized in their administration.

Effectiveness in Relationship to Quality of Management. The effectiveness of the operations of any organization is a function of the quality of its management. As with other operations, the effectiveness of a selection and classification program is a direct function of the effectiveness of management. In one investigation selective devices were studied in relation to this factor. The managements of a number of life insurance agencies were rated in terms of their effectiveness.[3] The validity of the selective devices used by managements segregated into three levels of

effectiveness were then determined. For those agencies wherein the management was judged to be better than average the validity coefficient of the selective devices was found to be .34. For those agencies in which the management was considered average in quality the validity was .25. For those companies in which the management was considered poor in quality the validity was only .03. Frequently, with ineffective management, selective devices are misused, the evaluation of worker performance is poorly done, and the selection results are distorted and misinterpreted. Satisfactory validity coefficients cannot be expected under such conditions.

Gains from Selection Programs. The gains from selection programs are manifold. Procedures developed for one purpose, *e.g.*, to increase production, may result in other types of improvements, *e.g.*, a reduction in accidents. Tests developed by the Air Force to predict success in early phases of training, were found to be useful in predicting flying safety at much later phases.[2] Various hidden benefits may result from well-developed procedures. In some instances it has been found that the mere installation of systematic selection procedures results in an improvement in the type of applicants.[12]

Not all of the gains from the use of systematic selection procedures are immediate; some are long term. In a study of substation operators a battery of tests administered at the time of hiring was found to have a validity coefficient of about .45 in predicting success 19 years later.[17] An investigation of the effectiveness of an intelligence test showed a correlation of .50 between scores earned at the time of hiring and level of job to which the workers were promoted some time later.[11]

A well-developed program of selection and placement, then, produces significant gains of many different kinds. In some cases the benefits are not immediately manifest and, in others, they cannot be directly measured in terms of dollars and cents. However, the gains are nonetheless real. The establishment of well-designed procedures for the selection and placement of personnel means more effective utilization of that most valuable "commodity," manpower.

REFERENCES

1. Brogden, H. E.: An approach to the problem of differential placement, *Psychometrica*, **11**, 139–154, 1946.
2. Crawford, M. P., *et al.*: "Psychological Research on Operational Training in the Continental Air Forces," Army Air Forces Aviation Psychology Program Research Reports, No. 16, 1947.
3. Ferguson, L. W.: Management quality and its effect on selection test validity, *Personnel Psychol.*, **4**, 141–150, 1951.

4. Ghiselli, E. E.: New ideas in industrial psychology, *J. Appl. Psychol.*, **35**, 229–235, 1951.
5. Harrell, W., and R. Faubion: Primary abilities and aviation maintenance courses, *Educ. Psychol. Measmt.*, **1**, 59–66, 1941.
6. Henry, F.: The theoretical efficiency of a test, *Research Quart.*, **18**, No. 2, 90–103, 1947.
7. Hull, C. L.: Variability in amount of different traits possessed by the individual, *J. Educ. Psychol.*, **18**, 97–104, 1927.
8. Hull, C. L.: "Aptitude Testing," World, 1927.
9. Jarrett, R. F.: Per cent increase in output of selected personnel as an index of test efficiency, *J. Appl. Psychol.*, **32**, 135–145, 1948.
10. Jenkins, J. G.: Validity for what?, *J. Consult. Psychol.*, **10**, 93–98, 1946.
11. Pond, M., and M. A. Bills: Intelligence and clerical jobs, *Personnel J.*, **12**, 41–56, 1933.
12. Stromberg, E. L.: Testing programs draw better applicants, *Personnel Psychol.*, **1**, 21–29, 1948.
13. Taylor, H. C., and J. T. Russell: The relationship of validity coefficients to the practical effectiveness of tests in selection: discussion and tables, *J. Appl. Psychol.*, **23**, 565–578, 1939.
14. Thorndike, E. L.: In "The Personnel System of the United States Army," Vol. II, "The Personnel Manual," 1919.
15. Thorndike, R. L.: "Personnel Selection," Wiley, 1949.
16. Thorndike, R. L.: The problem of classification of personnel, *Psychometrica*, **15**, 215–235, 1950.
17. Worbois, G. M.: Predicting long-range performance of substation operators, *J. Appl. Psychol.*, **35**, 15–19, 1951.

CHAPTER 6

The Interview and Personal-data Analysis

In current personnel practices the interview is the principal instrument used to obtain information about individuals. Private business and industrial organizations seldom hire an applicant for any position unless he is interviewed by some representative of management. The interview may be long and involved, consisting of several sessions between the applicant and one or more interviewers; or it may be a perfunctory affair lasting 5 minutes.

Great faith is placed by personnel officers in this method of assessing applicants. Although other selective procedures may be judged critically and rejected, one seldom finds the interview rejected, even after being carefully evaluated and found deficient. False feelings of satisfaction are engendered by its continued use. A strong self-assurance develops that one has the ability to assess human behavior subjectively. Objective analyses of interview results are then felt unnecessary, and frequently negative evidence discovered in such analyses is either rejected or misinterpreted.

In many aspects the interview is similar to the rating method. To a considerable extent generalizations and principles developed in connection with ratings will hold true for the interview. The types of errors made by raters are likely to be made by interviewers. Techniques found to improve ratings generally are found to improve the interview. Knowledge of problems and methods in the area of ratings, therefore, constitutes an important basis for understanding the interview.

KINDS OF INTERVIEW

The Unsystematic Interview. This refers to the casual, loosely organized interview which is in widespread use. In it no attempt is made to define specific areas of information to be explored. "To determine if the applicant is fit for the job" states the general objective. It is usually left to the judgment of the interviewer to introduce whatever details he feels necessary. Of course, this freedom given the interviewer is not

necessarily bad. It becomes unsound when the individual is not technically proficient in interviewing. As a rule, the individual chosen as an employment interviewer is simply one who has risen in the ranks of the personnel department. Although such a person is likely to be highly motivated toward his work and can honestly be described as being "interested in people," there too often is little in his background or training to justify his assignment to the task of interviewing. Beyond setting down rather general areas to be explored, he does little or nothing by way of planning the interview. The interview is conducted as a spur of the moment, inspirational affair. It must be remembered that the interviewer is doing a professional job and consequently should have the qualifications of a professional person.

It should not be concluded from the foregoing statements that the employment interview is unacceptable as a method for selecting employees. Rather it should be concluded that, placed in the hands of inadequately trained and inexperienced persons, and without proper procedural refinements, the method does not differentiate accurately between individuals with high potential and those with low potential. Under these conditions it is not an effective selection device.

The Systematic Interview. In the systematic interview there are two essential characteristics, namely, a careful systematic planning of the interview, and the use of technically proficient interviewers.[7] A systematic attack upon any problem can be expected to yield far better results than a haphazard casual procedure. Planning the interview, of course, is predicated upon knowledge of the job requirements. The interview is conducted to collect information of a precise kind about special types of worker performances from a particular individual. Only through a thorough knowledge of the requirements of the job to be filled and worker performances on this job can one understand the behaviors about which information is needed. Furthermore, job information points up the particular kinds of situations that must be explored in order to gain precise information for predictive purposes. Casual acquaintance with a job will not furnish the necessary basic understanding.

Systematic analysis results in clear statements of objectives, planned ways for obtaining the required information, and procedures for evaluating this information. It aids in making decisions concerning which areas can be adequately explored in the interview and which should be left for other types of selective procedures. It aids in correctly allotting interview time according to the importance of the areas and the difficulty anticipated in obtaining the necessary information. As a result a systematic interview provides a more organized approach and a more constant basis for the appraisal of different candidates.

Precision of results stems, in part, from the expertness with which the

interview is conducted. The nature and accuracy of the information divulged by an applicant will be a function of the confidence he has in the interviewer, the manner in which the questions are asked, the particular arrangement of the questions, and other similar factors. Such factors are directly associated with the technical proficiency of the interviewer. Not just anyone can be trained to be an expert in interviewing.

The Systematic versus the Standardized Interview. It is necessary to understand that there is a difference between a systematic interview and a standardized interview. In the latter, specific questions are set for the interviewer and he asks them of all applicants in exactly the same manner. The effect then is that the standardized interview achieves little more than can be obtained by means of a printed questionnaire filled out by the applicant. In the systematic interview, however, the number of questions exactly set will be few, if any. The systematic interview procedure provides guides concerning the general line of questioning, the specific questions to be asked being the responsibility of the interviewer and developing out of the interview situation. The skill of the interviewer therefore is of great importance in the systematic interview and is of little importance in the standardized interview. Getting away from the rigid structuring of the questioning found in the standardized interview allows for a more flexible approach and results in a procedure which can be adapted to each individual applicant. In addition it permits the applicant a better opportunity to present his case since he can discuss those points which seem to him to be most pertinent.[15]

EXPERIMENTAL STUDIES OF THE INTERVIEW

Paucity of Research Studies. Factual information concerning the interview as a systematic selective method is very fragmentary. To a considerable extent, research on the interview is still in the stage of formulating the basic problems. As a consequence the number of pertinent scientific studies of effective interview procedures is small. Indeed, the number of investigations on the basic problem of the validity of the interview is far less than would be considered sufficient to justify a final judgment on the adequacy of any other selective method. The interview, therefore, cannot simply be accepted as a thoroughly explored and empirically justified procedure for use in the selection of employees. Nevertheless many organizations that have refused to accept some other carefully evaluated selective procedure because the validity of that procedure appeared to be too low have continued to use the interview without any attempt whatsoever to evaluate its effectiveness. Generally the in-

terview is accepted and trusted on purely a priori grounds. Although such faith may be admired, it is necessary to ascertain whether it is justified.

The few studies that have been conducted on the effectiveness of the interview show a wide range of findings with respect to its reliability and validity.[23] While in some instances the method has been found to be very effective, in many instances the dependability is so low that it clearly is completely useless as a means for differentiating individuals in terms of their potentialities. It has been definitely established, however, that the interview does have possibilities as a means for appraising applicants. It is necessary to ascertain the particular conditions under which it is effective before it can be recommended for routine use as an instrument for personnel selection and classification.

Findings Concerning the Unsystematic Interview. Studies of the unsystematic interview generally have shown that its dependability is likely to be little better than chance. In a study now classical, Hollingworth compared the rankings made by 12 sales managers of 57 applicants interviewed by them.[5] The rankings of four of the applicants by the 12 interviewers, presented in Table 6-1, illustrates almost complete lack of

Table 6-1. Rankings of 4 out of 57 Applicants by Each of 12 Sales Managers

Applicants	Sales managers											
	A	B	C	D	E	F	G	H	I	J	K	L
I	53	10	6	21	16	9	20	2	26	28	1	57
II	33	46	6	56	26	32	12	38	9	22	22	23
III	54	41	33	19	28	48	8	10	26	8	19	56
IV	43	11	13	11	37	40	36	46	1	15	29	25

agreement among the sales managers. For example, applicant I was judged by sales manager K as being the best applicant in the group, and by sales manager L as being the worst. Additional discrepancies may be seen in the rankings assigned to the other cases. Kornhauser had a group of judges repeat their interviews with the same group of individuals.[9] The correlation between the two sets of interview ratings for the least consistent judge was .42, and for the most consistent judge, .78. These coefficients of correlation probably are too high to be used as true indices of reliability, since the judges were certainly influenced during the second interviews by the opinions they had formed in the first ones.

The results of the foregoing studies have been confirmed by other investigations.[23]

The general conclusion is that the unsystematic employment interview, as it is usually conducted by personnel officers, has a reliability far below what is necessary for the accurate measurement of individual differences. Since the validity of a technique is limited by its reliability, and usually is lower, it can also be said that the unsystematic interview has a validity which is too low for practical predictive purposes.

Findings Concerning the Systematic Interview. It is not surprising to find that in many instances the systematic interview has given quite adequate results. Thus Hovland and Wonderlic found that when the interview is systematically planned and placed on a professional plane, it may have very satisfactory reliability and validity.[6] Newman, Bobbitt, and Cameron report correlations ranging from .80 to .89 between the ratings of pairs of interviewers.[16] Other results obtained with the systematic interview bear out its validity. McMurry found validity coefficients ranging from .61 to .68 for various groups of industrial workers when a systematically planned interview was used.[13]

THE THREE COMPONENTS OF THE INTERVIEW

An interview consists of three components: the applicant, the interviewer, and the interview procedure. Understanding of the factors which condition the effectiveness of the interview can be gained from a study of these components and it then becomes possible to improve the dependability of the results.

The interview is a dynamic situation. The responses of the applicant are determined in part by the reactions of the interviewer. In turn, the procedures followed by the interviewer are organized with respect to the replies of the applicant. The role of the interviewer is not that of a bystander. He is both a stimulator and an inhibitor. He is an essential part of the interview method, a part which introduces change and, in turn, is itself changed as the interview proceeds.

THE APPLICANT

The Inconstancy of the Applicant. The applicant is a variable in the interview situation. Change in his mood is reflected in his behavior. Variations in environmental conditions produce marked differences in his ability to perform. The interview situation may provide stimulations that result in nonrepresentative performance and therefore inaccurate assessment of the applicant's abilities, interests and personality traits.

During the interview an attempt should be made to discover and ap-

praise the enduring characteristics of the applicant. It is the enduring qualities which are responsible for success on a job. Therefore it is necessary to differentiate between the enduring and the ephemeral aspects of the applicant's behavior. There are two procedures which aid in making this differentiation. One is to repeat within the same interview, but in modified form, questions dealing with the behavior under consideration. The other is to hold more than one interview, and thus have additional opportunity for discovering the stability of the behavior.

It is axiomatic in psychological measurements that repeated measures tend to improve the accuracy of estimation. Therefore, the results of the interview will be improved by affording a greater opportunity to learn the characteristic behavior of the applicant. This conclusion is supported by the findings that closer acquaintance with the individual being interviewed results in increased reliability of the interviewer's estimates.[17] Additional evidence comes from Kornhauser.[10] He found that some of the discrepancy between the judgments of two interviewers who rated the same persons was due to variation in the behaviors and attitudes of the interviewees on the occasions of the two interviews.

Adequate Motivation of the Applicant. One function of the interviewer is to see that the applicant presents a good case for himself. The point of view is often taken, however, that if an applicant is not sufficiently motivated to make a good case for himself, and as a consequence receives a low rating from the interviewer, he is not a desirable employee, regardless of his qualifications. Such a point of view is only justified if it can be shown that this particular expression of motivation is highly related to proficiency on the job. Although it has not been demonstrated that motivation in the interview is unrelated to job success, neither has it been demonstrated to be positively related to success.

Inadequate motivation in the interview may be due to a variety of factors; factors that actually may be unimportant as far as the evaluation of the individual is concerned. Thus an individual who is already employed and has a satisfactory job may simply "shop around" a bit. By not developing high motivation in such an applicant the interviewer may be overlooking a good prospect for employment. It should therefore be a primary function of the interviewer to motivate the applicant to give as much information about himself as he can, and to make the best possible case for himself that he is capable of presenting.

The Applicant as a Source of Information. The applicant does not always prove an accurate source of information. Sometimes this is attributable to the applicant, sometimes to the interviewer, and sometimes to the misuse of the interview method.

Limitations of the applicant's ability and personality may make it impossible for him to give an adequate picture of his potentialities. Some

individuals do not feel at ease in a face-to-face situation requiring them to talk. This is especially true when a marked difference in status exists between the applicant and the interviewer. It is necessary then to match the level of difficulty of the questioning to the verbal skills of the applicant. Emotional discomfort frequently accompanies an admission of error or failure by the applicant. If it is necessary to learn about the applicant's weaknesses or limitations, the questions should be of such a nature as to avoid stimulating an emotional reaction. If for any reason the applicant finds it difficult to express himself, it is not likely that he will give an accurate accounting of his experiences.

The accuracy of report of the applicant is associated with the skill of the interviewer. If the behavior of the interviewer causes the applicant to become suspicious of the use that might be made of his answers, he may be inclined to give only partial answers, or to distort the facts according to his own biases. Again, if the behavior of the interviewer elicits an emotional reaction in the applicant such as embarrassment, he may be inhibited from giving a full and accurate account. Under conditions such as those described great skill is demanded of the interviewer if the interview is to serve a useful purpose.

Frequently the interview method is misused. Attempts are made to gain information about certain kinds of attributes which are not demonstrable in the interview, cannot be readily evidenced in the interview, or are more quickly and accurately estimated by other means. As a simple example, arithmetic ability is better appraised by a test than by the applicant's memory of what grades he obtained in the subject in school. As will be seen in subsequent chapters, tests can give a variety of different types of information in a highly accurate manner, and certainly much more adequately than can be obtained by means of an interview. Similarly, information concerning an applicant's school grades can be most accurately determined from a school transcript. Of course, it is possible to learn about school grades in the interview. It should be remembered, however, that when a job is at stake there is a chance that the applicant may engage in a little stretching of his memory and the truth. There is no point in trying to evaluate the capabilities of an applicant by means of the interview when he is unable to present satisfactory evidence of them or is unable to demonstrate them.

THE INTERVIEWER

Individual Differences among Interviewers. Consideration should be given to the interviewer as a variable. It is obvious that the value of the results of an interview will be a direct function of the competence of the individual who conducts it. It has long been known that inter-

viewers differ among themselves in terms of the reliability and validity of their interview ratings.[23] Under the same set of conditions it has been found that the estimates of some interviewers agreed quite closely with those of some others, while being in complete disagreement with the estimates of still other interviewers. Similarly, interviewers vary widely in the validity of their estimates. Working under very similar conditions some interviewers are significantly better than others in predicting whether or not an applicant will be a successful employee.

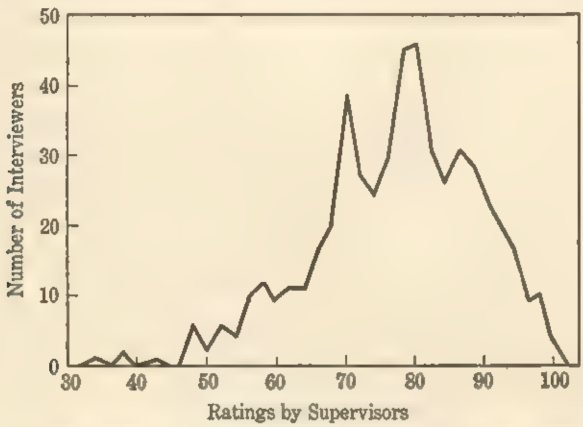


FIG. 6-1. Distribution of ratings of interviewers by their supervisors.

The general attitude about these differences appears to be that, although such discrepancies exist among interviewers, they are relatively insignificant. This view, however, needs further examination. The extent of differences among employment interviewers may be noted from a distribution of ratings made by supervisors of the proficiency of interviewers in the U.S. Employment Service.[18] The distribution of these proficiency ratings is shown in Fig. 6-1. In spite of the fact that these interviewers can be considered to be a fairly homogeneous group, since all of them had passed an appropriate employment examination, there is still wide variation among them in terms of their proficiency in the interviewing situation.

Selection of Interviewers. Recognizing that people vary markedly in their ability as employment interviewers, it is apparent that the interview would be made a very much better instrument if the interviewers were carefully selected on the basis of their ability to predict the future success of applicants. At present there are no tests or other standard devices by means of which good interviewers can be differentiated from poor ones. In the selection of their interviewers employment offices must depend primarily upon their knowledge of the interviewer's training and

past experience in the field of his particular industrial specialties. Of course, judgments about interviewing ability based on information about other specialties simply constitute good guesses.

Preselection, however, is not the only available approach to the problem. Procedures can be established to check on the accuracy of each interviewer's predictions. The validity of each interviewer's judgments should be determined in terms of the success of the placements for which he was responsible. Those interviewers who are found ineffective should be transferred to other jobs or should be given specialized training for their work. Obviously there is little point in retaining an individual as an interviewer if he is unable to judge with a fair degree of accuracy which applicants will be successful and which unsuccessful.

Training and Experience in Interviewing. Training and experience in interviewing improve an individual's reliability as an interviewer. Spielman and Burt found that, with increasing practice in interviewing, the reliability of interviewers' estimates increases.[17] Other investigators also have noted an increase in the consistency of the judgments of interviewers after training.[22] Thus the training of interviewers in the problems and techniques of the interview results in an increase in the dependability of the method. Since experience more or less amounts to informal training, experience—particularly if it is guided experience—should prove valuable.

While some thought has been given to the content of a training program for interviewers, it cannot be said that an outline of such a program has been clearly set.[20] On the basis of a purely logical analysis it would appear that the training should include at least four phases. The interviewer should be instructed on the nature and importance of individual differences as they are empirically demonstrated in the findings of scientific psychology. Secondly, he should be trained in the basic principles of assessment such as are discussed in certain chapters of this book. Thirdly, he should be given opportunity to develop a thorough knowledge concerning the specific requirements of the jobs for which he will appraise people. Finally, under guidance of professionally trained persons or experienced interviewers, he should perform trial interviews and have his performance critically evaluated.

Using More than One Interviewer. An incomplete picture and thus an invalid estimate may result from a single interview. This is particularly important when the job to be filled carries high-level responsibility and when the traits and qualities important for success are not accurately assessable by other selective devices. The fact that the correlation between the estimates of two judges is not perfect may indicate the operation of factors other than the unreliability of their judgments. In many cases two different interviewers will note different characteristics and

qualifications in the same applicant. Although each interviewer's estimates may be valid for the areas covered, the number of important qualifications he assesses may be incomplete. If the applicant is checked by a second interviewer, additional qualifications will probably be assessed. The pooled estimates of the two interviewers will then be superior to the estimates of either one inasmuch as a more complete picture of the applicant is obtained. Therefore, if the applicant is evaluated by more than one interviewer, the validity of the evaluation is likely to be higher.

THE INTERVIEW PROCEDURE

The Purpose of the Interview. The primary purpose of the employment interview is to obtain pertinent facts about the applicant's experience and qualifications as a basis for determining his suitability for a particular job. Of course, in addition, it serves to give the applicant some notion of the position he might obtain, its nature, its advantages and disadvantages, and its future possibilities. In many cases it is also used as a means for developing public good will by establishing a friendly relationship between the organization and the individual, regardless of whether or not he is hired. From the point of view of the interview as a selective device, the important consideration is that it is a means of obtaining information about a person in order that a proper evaluation may be made of his potentialities in connection with a particular job situation.

Securing Adequate Knowledge of Job Requirements. As with other aspects of the employment problem, knowledge of the job and worker specifications is essential for accurate interviewing. The interviewer should have at hand specific and detailed information concerning the nature of the work the individual will have to perform, and the qualifications necessary in order to perform the work satisfactorily.[23] Without such job information the interviewer will be unable to plan the interview so that all important areas of qualifications are explored and assessed. He will be unable to combine his appraisals in an optimal fashion to effect the best final rating of the applicant.

Planning the Areas of Qualifications to Be Explored. Having secured adequate knowledge of the job from the specifications, the interviewer is in a better position to plan systematically the course of the interview. He can outline with some degree of certainty the important areas to be explored. He can arrive at some conclusions concerning the types of questions most likely to be diagnostic of later success on the job. The more comprehensive the analysis that goes into this planning of the interview, the better the results are likely to be.

The importance of systematically planning the interview has been frequently stressed.[2] The value of planning lies in the fact that it provides the interviewer with certain guides relative to the kinds of information he should be seeking from the applicant. It gives him a constant background against which the qualifications of the applicant can be more accurately judged. Furthermore, it enables the interviewer to integrate all the information in a more meaningful manner in arriving at a final appraisal.[18]

The argument is often advanced that the use of such aids will eliminate informal discussion on the part of the applicant, and thus important information will not be revealed. Bingham and Moore, however, report that the use of such aids does not hinder but rather stimulates free and spontaneous discussion.[2] This would seem likely since both parties will have a clear understanding of just what constitutes the subject of discussion.

Securing Information before the Interview. In most worker assessments much information concerning the applicant can be obtained before he enters the interviewing situation. It follows from the proposition that employment interviewing is a highly specialized type of activity that those aspects of the worker's appraisal which can be performed by less specialized personnel or accomplished by other means should not be done by the interviewer. If the interviewer has at hand considerable pertinent information before he sees the applicant, he will be in a position not only to save time but also to plan more exactly the areas he feels need special attention for the particular applicant. Thus he may be able to specify his questions to a considerable extent. Before the interviewer begins his conversations with the applicant, therefore, it is desirable that he have whatever information can be made available on the qualifications and capabilities of the individual he is to appraise.

The Questioning Procedure. Although there have been a number of experimental investigations concerning the effects of various forms of questions upon the interviewee's answers, such investigations have failed to provide the employment interviewer with a series of principles to follow in the questioning. Questioning is still an art. For the most part, in order to prepare himself, the interviewer will have to depend upon systematic and logical analyses.

The extent of differences among interviewers in their participation in the questioning procedures has been found to be great. Uhrbrock noted the number of words spoken by the interviewer and by the applicant in typical employment interviews.[19] He found that, on the average, the interviewers did 60 per cent of the talking, ranging from one who did only 47 per cent to one who did 73 per cent. Obviously, if the interview is a fact-finding device the applicant should do most of the talking.

One important aim, which is a factor in determining the nature of the questioning procedure, is the necessity of securing from the applicant the answer he actually means to give. As was seen in connection with ratings, the same words may have different meanings for different individuals. It is essential that the interviewer understand the meanings contained in the answers of the applicant. The full meaning of each statement that the applicant makes should be secured. In case of doubt concerning a statement the question should be repeated in a different form, giving the applicant ample opportunity to make clear what he means. Complete agreement should be reached between the applicant and the interviewer on the meaning of all important points considered.

THE EVALUATION OF INTERVIEW DATA

Two Kinds of Final Evaluation. There are two opposing points of view in regard to the manner in which the information obtained in an interview should be evaluated in arriving at a final assessment of the applicant.[1] On the one hand, there is the view that the job should be broken down into its various specific aspects, and the applicant evaluated in terms of his qualifications in each of the aspects. The final appraisal should be the total or average of the ratings on all of the specific qualifications. The second point of view holds that such a breakdown is useful only in terms of the aid it gives the interviewer in covering all of the important areas. According to this view the final appraisal should not be based directly upon the ratings in each of the specific aspects of the job. The final assessment should be of the picture of the applicant as a whole. The former approach to the appraisal of the individual is analytic in nature and the latter is integrative. It is difficult to evaluate these two procedures. Certainly in terms of the evidence available at the present time it cannot be said with any high degree of certainty that one is generally superior to the other.

The Analytic Approach. The analytic point of view has one advantage in that it permits evaluation of specific phases of the applicant's qualifications that may be useful in later studies. When interviewers' ratings of particular qualifications are compared with measures of job proficiency obtained at a later time, it is ordinarily found that certain of the ratings are distinctly more valid than others. In Table 6-2 are given the validity coefficients of interviewers' ratings of applicants for positions in social-service work. It will be observed that in only one instance are the ratings on specific qualifications lower in validity than the assessment of the applicant's over-all fitness. Obviously, the total or average of the ratings on the specific qualifications should then have a higher validity than the rating of over-all fitness. To obtain the highest validity from a pooling

Table 6-2. Coefficients of Correlation between Interviewers' Evaluations of Applicants for Social-service Work and Later Obtained Measures of Job Success

Qualification	*Validity coefficient*
Appearance, poise and manner	.26
Comprehension	.39
Address and speech	.18
Education	.19
Training and experience	.19
Over-all fitness	.19

of the specific ratings, greater weight would be given to those qualifications found to be more predictive of later success.

The Integrative Approach. The argument against the analytic approach is that it analyzes the individual in a highly artificial manner which is likely to bear little or no resemblance to the total integrated person. The integrative approach allows the interviewer to take into consideration any compensatory effects that he might discover. For example, an applicant might not have quite the work history and past experience deemed necessary for a given job, but the interviewer, working on the basis of the integrative point of view, might hold that these deficiencies are unimportant since the individual is highly motivated and has exceptionally good work habits.

On the basis of the integrative approach the ratings on specific job qualifications would be expected to have low or negligible validity in the prediction of later success, but the validity of the assessment of over-all fitness would be expected to be high. In Table 6-3 are presented

Table 6-3. Coefficients of Correlation between Interviewers' Evaluations of Potential Executives and Later Obtained Measures of Job Success

Qualification	*Validity coefficient*
Educational background	.24
Occupational experience	−.10
Interest	.29
General intellectual ability	.00
Sociability	.34
Emotional stability	−.16
Initiative and drive	.08
Self-confidence	−.16
Over-all fitness	.71

data relative to the validities of certain qualifications of applicants in terms of their potentiality as executives, as evaluated by interviewers.

It will be observed that, for the most part, the validity coefficients of the ratings on specific qualifications are not greatly different from zero, but that the validity of the appraisal of over-all fitness is quite high. These results are directly opposite to those cited in Table 6-2 and discussed in the preceding topic.

Final Evaluation in Relation to the Qualifications of Interviewers. How can the discrepancy between the two sets of results of Tables 6-2 and 6-3 be reconciled? The probable answer can be found when the differences between the two sets of interviewers is taken into consideration. The individuals who interviewed the applicants for social-service work were either persons experienced and advanced in social-service work or persons who held administrative positions. They were not persons especially trained or experienced in interviewing or evaluating personnel. On the other hand, the interviewers who assessed the potential executives were professionally trained, with long experience in the appraisal and analysis of human capacities and traits. The accuracy of the assessment of over-all proficiency made by the first group of interviewers, those without specific training and experience in interviewing, was inadequate, although they were able to rate certain specific qualifications with somewhat greater accuracy. The trained and experienced interviewers, although not attempting to arrive at accurate estimates of specific qualifications, did provide an estimate of over-all proficiency which had high validity. It would seem therefore that when the interviewers do not have a specific background and training in techniques of psychological appraisal, it is better to have them evaluate applicants in terms of specific qualifications, and take as the final over-all rating the total or average of the ratings on the specific qualifications. For trained and experienced interviewers, however, it is likely that the appraisal of over-all fitness will have greater validity than the total or average of interview ratings on specific qualifications, since the skill of these interviewers will permit them to take into consideration the particular unique configuration of qualifications possessed by each applicant.

THE GROUP INTERVIEW AND SITUATION TESTS

Very frequently an examination of the educational and occupational histories of a candidate will reveal no pertinent evidence about certain qualifications that are important for the job. For example, if leadership ability is important the interviewer will attempt to uncover instances in the candidate's past history where he held positions of leadership, and will try to obtain some estimate of the quality of the candidate's performance in past leadership situations. If the candidate never held offices nor positions of responsibility over others, it might be concluded that the

applicant lacked leadership ability. From other information obtained about the applicant, however, it might appear that the circumstances were such that the applicant had little or no opportunity to participate in situations wherein he might have assumed the role of a leader. It is then necessary for the interviewer to draw conclusions concerning the leadership ability of the applicant on the basis of inadequate evidence.

The Interview as a Device for Measuring Personality-type Qualifications. One solution to the problem has been to present hypothetical situations to the candidate in the interview and to base evaluations of his qualifications on the nature of his responses. Thus questions like the following might be asked: "Suppose you were a foreman and the employees under you began talking too much. What would you do?" "If you were a sales manager, how would you deal with a salesman whose expense account was much too high?" Judiciously selected questions of this type frequently provide information which appears to be of some aid in making a judgment. The obvious difficulty, however, is that the kind of response being elicited is purely a verbal one. This is subject to errors of human judgment and bias. While an individual might say he would follow a particular course of action, when actually faced with such a situation or one like it, he might behave in quite a different way.

The Group Interview and Situation Tests. One procedure for overcoming this deficiency is to create a set of circumstances in the interview which compels the candidate to respond to a given set of conditions that will elicit the behavior to be evaluated. His behavior, including his verbal responses, is then appraised by observers.[11] Ratings are made on scales especially designed to indicate the extent to which the candidate manifests qualities deemed to be important.

A variety of situations have been utilized in an attempt to elicit pertinent types of behavior. In most instances candidates perform in a group, although in some instances they react alone. Some group situations may primarily involve argument and discussion. The observer presents a topic and the candidates, sitting around a table, offer various views and opinions. In this situation some individuals tend to initiate and lead the discussion while others participate very little. In some instances, the group is given a problem that they must solve working cooperatively.[8] The solution may call for only verbal responses, or it may require considerable physical activity. The group may be asked to discuss and provide a solution to a question such as "Should further federal taxes be obtained through increased assessments on personal income, or by means of a direct sales tax?" In other instances the problem set for the group may require the righting of an overturned boat, or the building of a small bridge across a creek. In all of these situations observers make notations concerning which candidates assume leadership, which get

upset, which want to go it alone, which show good judgment, and other similar qualities. Sometimes the behavior is recorded in objective terms such as the number of times aggressive or cooperative behavior is displayed.

Various names have been applied to the foregoing types of approach. When the situation merely involves group discussion it usually is termed the *group interview*. When the task involves physical activity in the solution of a problem the term *situation test* is likely to be applied. The common feature is that the performance or traits manifest in actual behavior are appraised by one or more observers.

Evaluation of the Group Interview and Situation Tests. Many advantages have been claimed for these procedures.[4] It is said that because judgments are made on the basis of actual behavior more accurate appraisals will result. Since a number of candidates are appraised at the same time, presumably sound bases are available for making comparisons between applicants. It might be expected that personality could be more accurately judged since an individual's traits are manifest in a broad social situation—one which is more realistic than is afforded by the ordinary interview. On the other hand, a number of deficiencies have been pointed out.[14] For one thing, the time for observation usually is no longer than that available for the ordinary interview. Again, since there are interactions among the persons involved, and since immediate environmental circumstances may place some applicants in a more advantageous position than others, the bases on which persons are compared may neither be comparable nor constant. Furthermore, while the appraisal situation is more realistic it also is quite specific, and therefore may not elicit behavior of a kind which is actually important on the job.

As with other selective devices, it must be empirically demonstrated that the behavior appraised is a valid predictor of subsequent job success. There have been relatively few evaluative studies of group interviews and situation tests. While there is considerable variation in the findings thus far, it does appear that the ratings obtained have some reliability.[12] Similarly, some studies of validity have demonstrated that adequate predictions of later performance are possible.[21] It is difficult, however, to compare the effectiveness of the group interview and situation tests with other techniques of appraisal. Certainly they are more cumbersome and difficult to administer than the systematic interview or standardized tests. As far as validity is concerned they seem to be superior to the unsystematic interview but not superior to a good, systematically developed interview. Psychological tests are very much easier to administer, and in the studies conducted thus far have been shown to possess equal or better validity. It is possible that future studies may

show that these new techniques are superior for certain types of occupations, *e.g.*, supervisory. At the present time there is insufficient information to make a final judgment.

PERSONAL-DATA ANALYSIS

Present Use of the Application Blank. Almost invariably in the employment process the applicant is called upon to report certain personal data. In addition to identifying information such as name and address, the application blank usually has items on educational and occupational history, marital status, etc. The evaluation of personal data as collected by the application blank has been even less thorough and critical than the evaluation of results obtained through use of the interview. In a survey made by the authors it was found that none of the personnel officers contacted had made empirical studies of the effectiveness of his application blank, yet none was willing to admit that his blank was not revealing information about later job success. In reporting on the manner in which they used their application blank as an aid in the selection of new employees, a large proportion of the personnel officers stated that they based their judgments primarily on the answers to a few "particularly revealing" items. None had studied their records, however, to verify empirically the predictive power of these items.

Inadequacies of the Application Blank. The successful use of personal-data items for selective purposes depends upon the realization of two conditions, namely, that the items are diagnostic of success on the job, and that the information given by the applicant is accurate. Such evidence as is available indicates that frequently neither of these conditions is achieved. If the predictions are valid, it follows that the items yield information concerning aptitude for the job. This, however, does not appear to occur frequently. Empirical studies of the content of application blanks reveal that many items have no relation to the particular job in question. Indeed, a number of investigators have found that for most of the items on the blanks there is little or no correlation between the answers and the criteria of later job success. Only about one out of five or ten items may be expected to show a forecasting effectiveness better than chance. Without pretesting the items it would be next to impossible to differentiate the diagnostic from the nondiagnostic ones. However, only a very small per cent of the application blanks in use are formed from pretested items.

The information given by the applicant will not necessarily be accurate. It is a common occurrence that pertinent aspects of home background, educational and occupational history, and other past experiences are forgotten or distorted with the passage of time. Certainly memory is se-

lective. At times the applicant may deliberately falsify. In addition he may misinterpret questions, and in some instances he may not even be qualified to answer them. If the information collected in the application blank is inadequate or distorted, valid predictions are impossible.

Systematic Construction of the Application Blank. The foregoing criticisms of the application blank should not be interpreted to mean that the method is completely invalid and therefore should be discarded. It has been empirically demonstrated that, at least for some job situations, the inadequacies of the blank can be eliminated by a systematic development of items for the particular job performance to be predicted. What is required is the same approach suggested for both the rating and interview methods. First, the areas of behavior related to the job are discovered. Next, items of biographical information in these and closely related areas are constructed. Then the items are administered to groups of applicants or of workers for whom some index of job success can be obtained. Lastly, empirical analyses are made to determine the degree to which each item is related to each job criterion.

It is by a procedure such as this that the subjective elements of the application blank are removed. It should be remembered that just because a biographical item logically appears to be related to job success is no justification for using that item without an empirical verification. Subjective judgments of this kind are valuable. They usually form the beginning of all selective devices. Without the follow-up empirical study, however, there is no assurance that a given item has any predictive power. With an empirical analysis the degree of validity becomes known. If the validity is low the item can be revised or discarded. If it is high the item can be retained with others of acceptable validity. The point is that the analysis furnishes a known dependability, that is, whether the validity is low or high, it is known and not presumed.

The empirical investigation of Cawl on the validity of biographical items illustrates the kinds of results that can be anticipated in these verification studies.[8] Some of his findings are reproduced in Table 6-4. The index of success was length of time on the job, and the responses of the applicants on each item were correlated with it. For clerical workers, only two of the ten items (numbers 1 and 6) show promising correlations with the criterion of success; for salesmen there are six items (numbers 1, 2, 3, 5, 6, and 9); and for accountants there are two items (numbers 3 and 6). These items are the immediately promising ones. The remaining items should be revised or discarded. It is perhaps worthwhile to note that an item which has predictive power for one job does not necessarily have predictive power for another.

Weighting Items on the Application Blank. The fact that, among items differentiating the better from the poorer applicants, some items

Table 6-4. Coefficients of Correlation between Answers on an Application Blank and Length of Time on the Job

Application-blank item	Clerks	Salesmen	Accountants
1. Length of time on last job	.22	.36	.07
2. Length of time on next to last job	.05	.37	.19
3. Average time on last 3 jobs	.18	.46	.27
4. Average salary on last 3 jobs	−.03	−.13	.07
5. No. of jobs previously held	−.17	−.22	.03
6. Age at time of application	.26	.31	.27
7. Height	.07	.02	.05
8. Weight	.10	−.07	.17
9. No. of dependents	.00	.20	.00
10. Amount of education	−.04	−.02	−.05

give a better prediction than others has led to the weighting of the item responses. The simplest procedure for obtaining weights is the so-called "horizontal per cent technique."[18] In this method workers who have been hired and for whom records are available are divided into two groups on the basis of some measure of job success. Typical groups would be the upper versus the lower third in production, workers who leave the job within a year versus those who remain on it for longer periods of time, or workers receiving high proficiency ratings versus workers receiving low ratings. For each item on the application blank the proportion of successful persons marking each alternative is calculated and taken as the weight for that particular alternative.

The hypothetical data of Table 6-5 illustrate the procedures by which weights are computed. These data are made up to represent the responses of 500 workers to three questions on an application blank. It is presumed that all applicants were hired, and after some time on the job they were divided into a better half and a poorer half on the basis of some index of job success. The figures in the second column represent the total number of workers who answered each alternative of each item. The figures in the third column represent the number of successful workers who answered each alternative of each item. The fourth column presents the frequencies of the third column expressed as per cents of the corresponding frequencies of the second column. Thus, at the time they were hired, 350 workers reported that they were married or widowed. Of these 195, or 56 per cent, turned out to be successful on the job. The figures in the third column may be interpreted as a very rough index of the chances for success of an applicant who responds according to a

Table 6-5. Calculations for Obtaining Weights of Responses to Items on an Application Blank

Items with alternative responses	Total number answering	Number successful workers answering	Per cent successful answering	Simplified weight
Marital status:				
Married or widowed	350	195	56	+1
Single	100	30	30	−2
Divorced or separated	50	25	50	0
Age:				
20 to 29	200	75	38	−1
30 to 39	150	100	67	+2
40 to 49	100	60	60	+1
50 and above	50	15	30	−2
Reason for leaving last job:				
Seeking better job	200	102	51	0
Job terminated	50	24	48	0
Fired	50	23	46	0
Job was unsatisfactory	200	101	51	0

given alternative. For example, for a worker who, at the time of applying, reports his age as between 40 and 49 years the chances are roughly six out of ten that he will turn out to be successful. The chances are roughly three out of ten for a worker who reports his age as 50 years or more.

The numbers in the fourth column could be used as weights of the answers, but for practical purposes the weights can be simplified. Since percentages under 50 show more chance of failure than success, they may be assigned negative values. Setting the range 46 per cent to 55 per cent as zero—50 per cent being the point of equal chance for success or failure—then 56 per cent to 65 per cent may be called +1; 66 per cent to 75 per cent, +2, etc., and 36 per cent to 45 per cent, −1; 26 per cent to 35 per cent, −2, etc. These simplified weights are shown in the fifth column. Weights are computed by this method for the alternatives of every item. In scoring an application blank the responses of an applicant to the item alternatives are assigned their appropriate weights. The sum of the weights of the alternatives he selects becomes his total score.

Reference to Table 6-5 will show that responses to the third item—reason for leaving last job—do not differentiate the superior from the inferior workers. Items like this one, with little or no differentiating power,

would be revised or discarded. For the remainder of the items responses would be scored in the manner indicated. An applicant who reports that he is married and is 35 years old would achieve a score of +3. One who reports that he is single and 25 years old would achieve a score of −3. The chances of the first applicant being successful are higher than those of the second.

In order to determine the over-all validity of the instrument, the total scores on the application blank are correlated with the index of job success. Scores developed from only two or three valid items will not provide stable predictive measures. Ordinarily ten or more valid items are necessary to achieve consistent predictive power.

The relationships between the responses to items on application blanks and job-proficiency measures are very sensitive to chance and sampling factors. Weights should then be developed on the basis of data collected from the responses of a large number of workers. In addition, constant follow-up studies should be made to detect any changes which might occur in the relationships between the items and the index of job success. Weights should be adjusted to correspond with any significant changes occurring in these relationships.

The validity of application blanks which are formed from weighted personal-data items has been demonstrated a number of times. A sum-

Table 6-6. Summary of Validity Coefficients Reported for Weighted Application Blanks for Various Occupations

Occupation	Validity coefficients for	
	Trainability	Job proficiency
General supervisors		.36
General clerks	.35	.34
Recording clerks	.60	.27
Sales clerks		.01
Salesmen		.42
Policemen	.35	
Vehicle operators		.48
Mechanical repairmen	.37	.48
Electrical workers	.32	
Structural workers	.29	
Processing workers		.58
Machining workers	.32	
Machine tenders		.38

mary of the validities of various blanks for different occupations is given in Table 6-6.* It will be observed from this table that, while there is considerable variation in the validity coefficients, for several occupations application blanks formed of weighted items are quite effective instruments for the evaluation of prospective workers.

Areas to Be Covered in the Application Blank. It is impossible to list the items that in general will insure valid results. As has been pointed out, items valuable for selective purposes for one job may be wholly inadequate for another job. It is possible, however, to list characteristics which usually merit investigation in any *trial* application blank. For a particular job, of course, many of these will be eliminated in the empirical analysis. Items should be tried out in the following areas:

Age
Marital status
Number of children
Number of other dependents
Height
Weight
Number of jobs previously held
Years of education
School grades
Business-college training
Average monthly earnings on jobs previously held
Average number of years on jobs previously held
Amount of life insurance
Amount of debts
Number of memberships in clubs
Types of jobs previously held
Reason for leaving last job

Another trend in the use of personal-history items is to seek information from the applicant concerning his vocational and avocational interests, reading habits, recreational activities, hobbies, etc. The use of such biographical items in the form of an inventory falls in the province of testing, a subject considered in the next two chapters.

REFERENCES

1. Allport, G. W.: "Personality: A Psychological Interpretation," Holt, 1937.
2. Bingham, W. V. D., and B. V. Moore: "How to Interview," Harper, 1941.
3. Cawl, F. R.: "A Method for Predicting Length of Service," Westbrook, 1926.
4. Fields, H.: The group interview test: its strength, *Pub. Personnel Rev.*, **11**, 139–146, 1950.
5. Hollingworth, H. L.: "Judging Human Character," Appleton-Century-Crofts, 1922.
6. Hovland, H. C., and J. T. Wonderlic: Prediction of industrial success from a standardized interview, *J. Appl. Psychol.*, **23**, 537–546, 1939.
7. Kephart, N. C.: "The Employment Interview in Industry," McGraw-Hill, 1952.
8. Killinger, G. G.: "The Psychobiological Program of the War Shipping Administration," Stanford University Press, 1947.

* The validity coefficients given in Table 6-6 are the averages of coefficients reported in a variety of investigations.

9. Kornhauser, A. W.: Reliability of average ratings, *J. Personnel Research,* **5**, 309–317, 1926.
10. Kornhauser, A. W.: A comparison of raters, *J. Personnel Research,* **5**, 338–344, 1927.
11. MacKinnon, D. W.: "Assessment of Men," Rinehart, 1948.
12. Mandell, M. M.: Validation of group oral performance test, *Personnel Psychol.,* **3**, 179–185, 1950.
13. McMurry, R. N.: Validating the patterned interview, *Personnel,* **23**, 263–272, 1947.
14. Meyer, C. A.: The group interview test: its weakness, *Pub. Personnel Rev.,* **11**, 147–154, 1950.
15. Moyer, N. A.: Non-directive employment interviewing, *Personnel,* **25**, 377–396, 1948.
16. Newman, S. H., J. M. Bobbitt, and D. C. Cameron: The reliability of the interview method in an officer-candidate evaluation program, *Am. Psychol.,* **1**, 103–109, 1946.
17. Spielman, W., and C. Burt: The estimation of character qualities in vocational guidance, *Ind. Fatigue Research Bd.,* No. 33, 1926.
18. Stead, W. H., and C. L. Shartle: "Occupational Counseling Techniques," American Book, 1940.
19. Uhrbrock, R. S.: Analysis of employment interviews, *Personnel J.,* **12**, 98–101, 1933.
20. Uhrbrock, R. S.: The personal interview, *Personnel Psychol.,* **1**, 273–302, 1948.
21. Vernon, P. E.: The validation of civil service selection board procedures, *Occupational Psychol.,* **24**, 75–95, 1950.
22. Viteles, M. S.: "Industrial Psychology," Norton, 1932.
23. Wagner, R.: The employment interview: a critical review, *Personnel Psychol.,* **2**, 17–46, 1949.

CHAPTER 7

Characteristics and Development of Psychological Tests

Of all of the instruments devised by psychologists for appraising human abilities and traits, tests are clearly the most refined. Tests possess many advantages over other selective and evaluative devices. They are more amenable to objective investigation and hence more amenable to analysis and validation. This means that the amount of knowledge concerning them is greater. Administratively tests are much superior to interviews and rating procedures since their administration requires less highly trained personnel. Unlike the results from interviews and ratings, the results from tests usually do not directly reflect the personal qualifications of the administrator. A test can be administered by anyone with adequate background and training and still give comparable results. Because tests, in the main, are objective, they are impartial. The biases of the test administrator affect the evaluation of an individual to a much smaller extent than do the biases of an interviewer or a rater. Finally, tests yield quantitative descriptions of an individual's capacities and characteristics. Such quantification is a most important consideration in procedures used for differentiating individuals.

USE OF TESTS IN BUSINESS AND INDUSTRY

Psychological tests have made very significant contributions to the solution of problems of worker selection, placement, and adjustment. This is attested to by the ever-increasing demand for tests being made by business and industry, and by the growing number of fields of work for which tests are being devised. The successful use of tests by various government services and private industries during the war years added a marked impetus to the testing movement. The outlook for the future is one of greater test development and application in the selection, classification, and adjustment of workers.

In spite of the important advantages of tests, many organizations do

not utilize them, preferring to continue with established procedures such as the interview. In part this reaction arises from ignorance concerning the usefulness of tests, and in part from a belief that the interview, even though it is unsystematic, must necessarily be better because of its long tradition. Neither of these is a sound reason. An organization that refuses to utilize tests simply on the basis of a lack of knowledge or an unfounded bias, rather than on the basis of empirical fact is not keeping abreast of the advancements being made in personnel practices.

Tests have many uses in an integrated personnel program. Obviously test scores are useful in the selection and classification of workers. They also contribute to training programs by indicating those individuals who can profit most from training, by pointing up those phases of job proficiency wherein the individual is deficient and hence in need of retraining, and by measuring progress resulting from training. A number of studies have shown that tests are helpful in indicating potentiality for promotion. Lastly, tests provide information useful in programs for the upgrading of workers.

CHARACTERISTICS OF TESTS

Tests as Predictors. When tests are used for personnel selection and placement, the test scores serve as predictors of some index of job success. In selection the test scores are obtained before hiring the worker. They serve as a short cut for getting information on the probable success of the worker before he actually is hired. Before test scores are used in this way, however, they must be checked against some measure of job success. An index of job success is therefore needed as a means for determining the predictive effectiveness of the test scores.

This process of determining the degree of effectiveness is called validation. A test that has been shown to predict successfully an index of job proficiency is said to be a valid test. Some adequate measure of validity is absolutely necessary before the value of a test can really be known and before the scores on a test can be said to have any meaning as predictors of job success. Actually the significance and value of a test can be accurately obtained only when the test is checked against an adequate, independently determined index of job proficiency. Of course, the picture is somewhat complex, because different tests will vary in terms of their effectiveness in predicting a particular index of job proficiency, and the same test will vary in terms of the effectiveness with which it will predict different indices.

The use of unverified tests, whether through innocence or intent, cannot be condoned. Frequently, unfounded claims are made for tests, and the test scores are interpreted with a certainty and finality not supported

by the statistical facts. Occasionally a naïve faith is found which ascribes almost an attribute of magic to the test scores.

Referring Tests to Definite Indices of Job Success. In order to validate a test the test scores must be referred to specific measures of job proficiency. In the chapter on general principles of selection and classification, it was pointed out that there are many ways of evaluating a worker's proficiency on a job. This means that a test can be a valid predictor of one index of success and not of another. Similarly, a test may be very useful in predicting success on one job and be of little or no value in the selection of workers for a different job. A test, then, is not inherently valid or invalid; rather, it is effective for those particular purposes for which the empirical evidence justifies its use.

In describing the validity of a test it is necessary to describe the index of job success against which it was validated. If the index itself is not valid, then a test predicting it is not useful, even though it accurately measures the index.[2] If the index used as a measure of proficiency does not include all of the phases of job success, a test correlating with it can be considered to predict only those phases of the job measured by the index. Furthermore, a test is not to be considered valid simply because it contains elements which logically appear to be similar to certain elements found in the performance of the job. For example, if a test is known to measure some psychological ability, such as ability to work with mechanical relations, and certain mechanical performances are required in the performance of the job, the test still cannot be considered valid until the scores have been checked against some index of job success.

Tests as Samples. As indicated earlier, tests are short-cut methods for the prediction of job success. They are used as a substitute for actually placing an individual on the job and seeing how well he performs. Therefore, as with any short-cut method, a test does not measure in its entirety the variable to be assessed but merely samples the variable. For example, complete information concerning the quality of cotton in a given bale is best obtained by studying every fiber in the bale. Since this is impractical, a sample of the fibers is studied. In the assessment of human abilities, the same principle operates in the use of any psychological measuring instrument. Samples of a variable must always be considered as just what they are, merely samples. Of necessity there will be errors in the process of sampling, and it is important to understand the causes of these errors. The three ways in which the sampling process operates in the use of tests will now be considered.

Tests as Sample Stimuli. First, a test situation is an artificially organized and controlled situation devised for eliciting a given type of response from the testee, wherein means are provided to record both the quality and quantity of the behavior for later evaluation. A given test is a sample

of many possible stimulating situations that might be used. It plays the role of an exciting cause to initiate behavior. It does not determine or cause the responses, but merely acts to elicit them. The fundamental causes, or the factors determining the responses, are to be looked for in the native dispositions and past experiences of the individual. The test, however, narrows down the range of dispositions and experiences allowed to function.

It should be apparent that the test as an exciting stimulus will not necessarily have the same exciting effect on every testee; *i.e.*, it will not be equally effective in getting all testees to respond in terms of the same kinds of dispositions and experiences. Whether the ability or trait being measured is the same for every individual is dependent upon whether the stimulating items have the same meanings for every testee. Certainly, ambiguous items that can be interpreted in several ways cannot be considered to function as constant stimuli. Regardless of how much of a given ability is possessed by two individuals, a test may prove very effective with one and ineffective with the other in stimulating them to manifest this ability through response.

Tests as Samples of an Ability. Secondly, the responses to the test are to be considered as merely one sample of the many ways in which an ability can be manifested. The particular combination of native dispositions and past experiences called into play are determined primarily by the test. Whether the responses to the test represent the ability desired can be determined only through validation studies and should not be accepted on a priori grounds. They become an accurate sample to the degree that they represent in quality and quantity the type of manifestation later to be found on the job.

Tests as Samples of the Individual's Ability. Thirdly, the responses to a test made by an individual on a given occasion are to be considered only one sampling of his ability. An individual given the same test on two occasions will not make exactly the same score, and sometimes the discrepancy between the two scores may be very large. If measurement is to be accurate it is necessary to obtain responses which are characteristic of the individual. This is accomplished by including in each test a large number of items eliciting expressions of the same type of ability, and by utilizing several tests rather than merely one test.

At best, however, measurement merely samples the individual's ability. For a short period of time, in a test situation he responds to a set of problems or questions. These responses are considered characteristic of his behavior, and the generalization is made that the individual would behave in the same way in all situations of a similar nature. Such a generalization is accurate only to the degree that the sample responses are representative of the individual's ability. It must always be kept in mind

that every sample is an approximation, and that any generalization must be applied within this limitation.

Implications of Tests as Sampling Devices. From the foregoing discussion it should be clear that any interpretation or meaning given to a set of test scores must be conditioned upon the accuracy with which the test samples the same ability in different individuals, the degree to which it samples the ability that one desires to measure, and the extent to which it accurately samples each individual's ability. The attainment of adequate sampling in each of these ways does not present insurmountable difficulties. An understanding that tests are sampling devices is necessary in order to counteract the blind faith in test scores frequently aroused in those persons not familiar with the statistical logic underlying measurement procedures. Individuals desiring to use tests should devote sufficient time to their study to gain the realization that, like every tool yet invented, tests lose their precision and usefulness when utilized without critical understanding.

POTENTIALITY VERSUS PROFICIENCY

As stated above, the determinants of an individual's test responses are found in his native dispositions and past experiences. In the psychological measurements involved in personnel selection and placement there is no need to distinguish between innate and acquired factors. However, certain fundamental distinctions in the purposes of measurement merit further consideration.

Meaning of Terms Used to Designate the Testee's Performance. One purpose of measurement in industry is to describe what the worker can do at the time of testing, the emphasis being on the evaluation of responses which would be immediately useful on a job. A second purpose is to estimate from the worker's responses the power he has to accomplish future growth and development, the emphasis being on potentiality for later accomplishment. Several terms have been used to distinguish these purposes.

The word *ability* is used in a very broad sense to mean power to perform. It does not imply that the power is actual or potential, and makes no distinction between native dispositions and acquired experience. If the purpose of the measurement is to select a worker who is qualified to perform immediately in a given situation, the word *proficiency* is used. Proficiency refers to the worker's present accomplishments or achievements—what he actually can do at the present time. On the other hand, the word *potentiality* refers to dormant ability which awaits development. The words *aptitude* and *capacity* are also used in the sense

of potentiality or possible future response. No distinction is made as to whether the potential is conditioned upon native or acquired factors.

Examples of Proficiency and Aptitude Test Items. Table 7-1 contains several items from an aptitude test and several items from a proficiency test, both used for evaluating streetcar conductors. Although scores on certain of the items in the aptitude test will be influenced to some extent

Table 7-1. Examples of Items from an Aptitude Test and a Proficiency Test, Both of Which Were Used in Evaluating Streetcar Conductors

Items from an aptitude test for streetcar conductors

Put a plus (+) between any of the following pairs of numbers that are the same and a circle (o) between any that are different.

648396_____648396
960584_____960684
396759_____397659
104782_____104782

Add or subtract the following numbers as indicated.

948693	105846	574106	473017
−938610	+285901	−396907	+185739

Put a check mark before the best answer in each of the following questions.

The best way to face when jumping off a moving streetcar is to face:

_____the front of the car
_____the sidewalk
_____the rear of the car

Items from a proficiency test for streetcar conductors

Put a check mark before the correct answer in each of the following questions.

What is the position of the farebox on the car?

_____In pulling out of the carhouse, put it on the front end of the car.
_____The farebox must be kept on the rear end of the car at all times.
_____When leaving the car to get change, the conductor takes the farebox with him.
_____When changing ends, the conductor puts the farebox in the sand box.

Where are number plates placed?

_____Run number plates are placed on the side windows of cars.
_____Run number plates are placed on the conductor's end only.
_____Run number plates are displayed on both ends of the car.
_____Conductor and motorman hang run number plates on left lapel of coat.

In punching the time on the transfers for "Owl" cars:

_____Allow 3 hours from end of line.
_____Allow 15 minutes after midnight.
_____Allow 1 hour.
_____After 12 P.M. punch "emergency."

by the amount of formal schooling and general experience of the individual, beyond a relatively low minimum, additional amounts of education and experience probably will not have any great effect. The important point here is that previous experience as a streetcar conductor is not necessary in order to obtain a high score on the aptitude test. On the other hand, in order to obtain a satisfactory score on the proficiency test, experience as a conductor is necessary.

Difficulty of Measuring Proficiency and Potentiality. The measurement of proficiency is a somewhat simpler problem than the measurement of potentiality. To measure proficiency in a given task, a test can be constructed in which the responses required of the testee very closely simulate the responses used by the worker in performing the task. The more closely the test situation reproduces the elements of the task, the higher the validity of measurement is likely to be.

On the other hand, the evaluation of potential ability requires the measurement of some present performance that, in the nature of the case, cannot be the same as the subsequent ability. The ability to be estimated being a potential, it cannot be expressed in the test situation in the form that subsequently is to be predicted. The test performance must be so related to the potential ability, however, that it can be used as a predictor of the desired future performance. Other abilities that have been developed and that are expressible in present performance must be found which are closely related to the potential ability. It has frequently been demonstrated that it is possible to use the proficiency in one area as an index of the most probable amount of potentiality in another area. Thus, high school achievement is used as an index of potentiality for college work. Great difficulty is often encountered in discovering proficiencies which are closely related to the potentiality to be estimated.

CLASSIFICATION OF TESTS

In view of the fact that psychologists have been devising tests for nearly half a century it is not surprising that many different kinds have been evolved. To try to describe all of the different kinds of tests now available would be an insurmountable task. Perhaps the best way of obtaining knowledge about the different kinds is to classify them in terms of several large descriptive categories. In the following pages the more important of these categories will be discussed.

Classification in Terms of the Behavior Characteristic Being Measured. One major classification of tests is in terms of the psychological nature of the characteristic which the test is primarily designed to measure. It may be intended that the test measure ability of one kind or another, or it may be intended that the test measure some aspect of personality,

temperament, or interest. Personality, temperament, and interest do not readily lend themselves to measurement by tests, so other means, especially questionnaires and inventories, are more frequently employed. Because of the special nature of these instruments they will be given separate treatment later in the chapter.

Of course, it cannot be said that any test or inventory measures only ability or only personality. Certainly both kinds of characteristics contribute to the scores on any test or any inventory. Thus on a test designed to measure ability to follow complex directions, a worker who is low in perseverance will obtain a low score. Similarly, on a questionnaire designed to measure sociability, a worker who has only fifth-grade reading ability, but who nevertheless is sociable, will perform in an inadequate manner.

Classification in Terms of the Objective of the Testing. A test of ability may have either one of two objectives as follows: to predict how well a person will do on a job for which he has had no training, especially training in the specific procedures and work techniques utilized on the job, or to estimate how well a person who has had such training will do on the job. The difference here is between potentiality and proficiency, which has already been discussed in detail.

Tests of potentiality or aptitude may be further classified. Some tests of this kind purport to measure a general aptitude and others specific aptitudes. For example, there are tests of general mechanical aptitude which presumably measure future success in any kind of job involving mechanical ability; and tests of general clerical aptitude which presumably predict success in any kind of a clerical job. On the other hand, a specific aptitude test is designed to predict success in a specific job, such as a test for garage mechanics or a test for file clerks. Obviously, not all aptitude tests are either general or specific, but vary through many increments between these two extremes. The following jobs require tests of increasing generality: comptometer operator, business-machine operator, clerical worker primarily concerned with numerical operations, and, finally, general clerical worker.

General aptitude tests are important aids in the classification of workers. In those industrial organizations where an attempt is made to classify and optimally place each individual, it is important to obtain as much knowledge as possible about the worker's abilities. If an organization has a variety of jobs, many of which are either mechanical or clerical in nature, a test of general mechanical aptitude and one of general clerical aptitude will greatly aid in determining the placement of workers in the organization.

A further differentiation of proficiency tests may also be made. Pro-

ficiency tests are classified into those primarily measuring job knowledge or information and those primarily measuring job skill. In the first instance, proficiency is measured by the amount of information about the job that the worker possesses. In the second instance, it is measured by the amount of skill the worker displays in carrying out the job. The early proficiency tests in industry were developed for the skilled trades, and so came to be called *trade tests,* even though some were devised for occupations which were not of a mechanical nature. Although most trade tests are, perhaps, measures of job information, many of them are truly tests of skill. Consequently, the term *trade test* may be used to refer to either type of test.

Work-sample tests are examples of tests measuring job skill directly. In tests of this kind the worker is required to perform the same tasks that the job requires. Work-sample tests have been discussed in detail in an earlier chapter. It should be pointed out that there is an important difference between a job-skill test and a simple demonstration of proficiency. In a job-skill test the testing procedure is standardized, the problem assigned the testee is the same for all individuals, and performance is evaluated by comparison with established norms.

Table 7-2. Distributions of Scores of Three Groups of Workers Varying in Experience in Bricklaying on a Job-information Test for Bricklayers

Score on test	Expert bricklayers	Apprentices and helpers	Workers in related jobs
15	xx		
14	xxxxxxxx		
13	xxxxxxxxxxxxxxxxx		
12	xxxxxxxxxxxxxxxxxxxxx		
11	xxxxxxxx	x	
10	xxxxx	xx	
9	xx		x
8	xx	x	
7		xx	
6		x	x
5		xxxxxx	
4		xxxxx	x
3		xx	xxx
2		xx	xx
1		xxx	xxxxxxxxxxxx
0			xxxxxxxxxxxxxxx
Number of cases	65	25	35

Although job-information tests classify workers according to job knowledge, owing to the fact that the questions differentiate individuals varying in experience, these tests also classify workers according to job skill. In Table 7-2 are shown the distributions of scores on a job-information test for bricklayers made by workers varying in amount of experience on the job of bricklaying.[6] The validity of the test is clearly demonstrated by the fact that groups with more experience on the job earn higher scores on the test than those with less experience.[10] Further evidence that job-information tests measure skill comes from the fact that the test scores are usually related to proficiency on the job. In Table 7-3 are provided

Table 7-3. Coefficients of Correlation between Performance on Several Job-information Tests and Supervisors' Ratings of Proficiency

Test	Number of cases	Coefficient of correlation
Magnetos	24	.62
Warehouse (women)	150	.43
Sheet-metal repair (women)	90	.36
Sheet-metal repair (helpers)	82	.34
Sheet-metal repair (men)	63	.32
Service mechanics	130	.31
Sheet-metal repair (juniors and journeymen)	216	.18
Warehouse	160	−.03

the validity coefficients of several job-information tests used in an Army depot.[7] Validity is measured by the correlation between the test scores and ratings of proficiency by supervisors. It will be noted that in six of the eight studies substantial correlations were found between the test scores and proficiency on the job.

Classification in Terms of the Nature of the Response. Tests are also classified according to the manner of response required of the testee. A broad distinction is made between verbal tests and performance tests. In verbal tests the response is either written or oral. In performance tests the individual operates apparatus or machines, assembles objects, or performs other manipulations.

Classroom examinations are examples of written tests, since the response to the questions is accomplished by means of writing something down. Job-information tests may be orally administered in connection with the employment interview. The questions on the test are read by

the interviewer and the applicant gives his answers by means of the spoken word. Such a test is an oral trade test. Examples of performance tests are tests of finger dexterity in which the task may be to insert pegs into holes, and tests of mechanical aptitude in which the task may be to assemble simple objects.

In most instances, verbal tests have an advantage over performance tests in situations where either can be used. Verbal tests are less expensive and more easily administered. When the responses are written, the tests can be administered to a large number of persons simultaneously. The major advantage of performance tests is that they afford an opportunity to observe the individual as he solves the problems presented to him. Often from such observation information can be obtained relative to such factors as mode of attacking a problem, persistence, willingness to try different approaches, and other characteristics associated with work habits.

Classification in Terms of Mode of Administration. Tests can be classified on the basis of whether they are individual or group tests. Some tests, such as those requiring oral responses, can be administered to only one individual at a time. A good example is the Stanford-Binet Intelligence Examination. Other tests, such as those requiring written responses, can be administered to large numbers of individuals at the same time. Most civil service examinations include one or more group tests. From the point of view of testing time, the advantage of the group over the individual test is apparent. Furthermore, it ordinarily requires a less skillful individual to administer a group test. However, not all of the advantage lies with the group test. As pointed out above, sometimes when the individual test is used the personnel officer has an opportunity to observe the testee's performance at firsthand and thus is better able to make ratings of the individual on other traits, should this be desired.

From the point of view of reliability and validity, neither test has an advantage over the other as far as the industrial situation is concerned. It sometimes is held that the confusion and social stimulation found in group testing markedly affect test performance and thus reliability and validity. There seems to be little empirical evidence for this point of view. Group tests have been devised which are fully as adequate in respect to reliability and validity as are the best individual tests.

Classification in Terms of Testing Time Allotted. In some tests the individual is permitted to work on the items for only a specified time interval regardless of the number of items he is able to complete, or he is given a certain number of items and required to answer them as rapidly as he can. In other tests he is given a generous period of time so he can work on every item without being hurried. The former type of test is called a speed test and the latter a power test.

The first type of speed test described above is called a time-limit test. The second kind is called a work-limit test. In the former the individual is scored on the basis of the number of items he completes in the given time interval, and in the latter he is scored in terms of the time he takes to complete the given number of items. For example, in a speed test of typewriting, either the number of words typed per minute or the number of minutes required to type a certain amount of material can be used. In the time-limit method more items are included in the test than the best individual is able to complete within the time permitted. The time-limit method is the more widely used because of its convenience, the work-limit method being impractical when a group of individuals is to be tested.

In a power test the individual is given a liberal amount of working time. All testees have enough time to work on all of the items. Usually the difficulty of the items increases gradually from the beginning to the end of the test. The individual may then find that he reaches a point beyond which he does not know the answers. Because of the generous time allotted he seldom attributes his failure to the shortness of the time period, and usually acknowledges that it reflects the limit of his own ability.

Although speed tests are advantageous from the point of view of administration, they are disadvantageous in other respects. It has been found that older adults do not perform up to par on tests depending primarily upon speed of response.[4] Especially with older workers speed tests may reflect a level of ability much lower than that truly characteristic of the individual. In cases of emergency, as during World War II, when it is necessary to draw upon older individuals to fill labor shortages, the use of speed tests gives an inaccurate picture of the abilities of the individual. In such instances power tests give measures that are far more representative.

Classification in Terms of the Standardization of the Test. Another general basis for classifying tests is the amount of information furnished about a test which can be used for interpreting the meaning of the scores. The terms commonly used for distinguishing tests on this basis are *standardized* and *nonstandardized*. However, there is no clear-cut division of tests with regard to this characteristic as is suggested by the use of these two terms. Generally speaking, the standardization of a test involves setting up conditions for administering, scoring, and interpreting which will produce the desired level of validity. The constructor of the test is responsible for developing these conditions. It is obvious that standardization exists in degrees or amounts, and that the more adequate the conditions set up by the test constructor the better the standardization. Classifying tests as either standardized or nonstandardized ignores

the continuousness of this characteristic. Further attention will be given to test standardization in subsequent topics.

CHARACTERISTICS OF A GOOD TEST

Thus far in the discussion, several characteristics desirable in tests have been mentioned. It will be profitable to reconsider these briefly and to discuss other characteristics to be looked for in any test contemplated for practical utilization. It should be stated at the outset that every test already published cannot be expected to measure up to the standards to be presented.[3] Neither will the initial attempts of a test research program in a given business or industry always result in instruments which meet all of the standards. However, there is ample proof from past achievements in testing, not only in the academic field in which tests have been developed to a very high level of effectiveness, but also in the fields of government and business and industry, that tests of high predictive value can be constructed if adequately trained personnel and time and funds for needed fundamental research are available.

Validity. There are three approaches to a problem requiring the use of tests, namely, to use an available test purportedly designed for the purpose in hand, to adapt for this purpose a test constructed for some other purpose, or to construct a completely new test. In each of these approaches the first characteristic to look for or to achieve is that of high validity. Inasmuch as the forecasting of a worker's subsequent performance from his test score is involved in the use of nearly every test, validity is synonymous with effectiveness in prediction. As was indicated in a previous section, predictive value must be stated in terms of a definite criterion measure of job proficiency. Even a more specific statement can be made. In employee selection a test score should successfully predict an applicant's performance on a specific job, as measured by a particular index of job success, after a definite amount of training of a specific type has been given.

Whether or not a test should be finally adopted for use is determined in part by the effectiveness of the methods already employed. The level of predictive effectiveness to be accepted for a test should be determined for each specific situation in which the test is to be utilized. If a testing device does not significantly improve prediction of future job proficiency above that which current procedures achieve, it usually does not merit adoption. Whether a test is judged to have sufficient validity or not is conditioned upon the success attained by the methods the test is designed to supplement or supplant.

It is difficult to determine whether a test will effect an improvement over present procedures without actually trying it out. Few studies have

been conducted in industrial situations to determine the success of current non-test procedures so that a quantitative statement of the predictive value of these procedures is seldom available. The one common standard against which all tests can be evaluated is that of chance. It is possible at the end of a validation study to state what improvement over chance success is obtainable from using the test. As discussed in Chap. 5, this expected improvement is conditioned upon the following factors: the percentage of applicants making good on the job when no testing program is used, the heterogeneity of workers in terms of job proficiency, the available supply of labor from which new employees are to be chosen, and the degree of correlation between scores on the test and subsequent success on the job.

Before leaving the topic of validity, brief mention should be made of what is known as *face validity*. Face validity refers to the apparent validity that a test has because it demands from the testee a response that rationally or logically appears to be related to the performance to be predicted.[5] For example, if the task is to predict ability to drive a truck, and the test consists of an apparatus simulating the driving compartment of an automobile including a steering wheel, clutch, and brake which are used in responding to various signals, then the test is said to have face validity. Face validity is useful on several counts but, primarily, as a means of enlisting favorable attitude. Testees who know the requirements of the job for which they are being considered are quick to detect any resemblances between the test situation and the job situation, and most testees are favorably disposed toward any test in which they find an apparent similarity. In the end, however, the validity must be determined and expressed in terms of predictive effectiveness.

Reliability. Another characteristic every test should have is high reliability. Reliability refers to the consistency with which the test operates in repeated trials. Any measurement, regardless of the precision of the measuring instrument, will always be subject to a certain amount of error. A person measuring the length of a 10-ft. table to the nearest 0.1 in. will not get identical lengths each time he performs the measurement. If a person is given two forms of the same test on separate occasions, the two scores will not be the same. His performance on the tests is affected by his physical condition, his incentive and attitude, distractions present during the testing, and other similar factors. Such factors would not be exactly the same on the two testing occasions; consequently some difference in the individual's two scores is to be expected.

Reliability should not be confused with validity. The fact that a measurement is reliable does not mean that the device is accurately measuring that which it is expected to measure. An aptitude test may give consistent results in the sense that individuals upon repeating the

test may get about the same scores, but this consistency in no way indicates the extent to which the test actually measures the aptitude intended. The latter is the problem of validity or predictive effectiveness. If the results are consistent, however, it is assumed that what is being measured is being measured reliably or accurately.

Reliability is of considerable importance in testing. The value of any score in psychological measurement lies in the fact that it forms a basis by which an individual's performance can be evaluated in terms of the performances of other individuals. This evaluation may take the form of ordering the scores to compare the individuals in terms of their ranks, or more exact and refined methods may be used for indicating the amounts of the differences between the scores of the various individuals. Regardless of the technique used, test scores form the basis for drawing inferences and making administrative decisions about individuals who obtain different scores, or about groups with different mean performances. It is apparent that for these inferences and decisions to be of any value the measurements must be reliable. To attach any importance to a test score, it must be known that when individual A gets a score higher than individual B the difference between the two individuals would be similar if the test were given again under similar conditions.

Suitability of the Test for the Group under Consideration. In the selection or design of a test, it is important to bear in mind the nature of the group on which it will be employed. The suitability of a test may be indicated in a variety of ways. One simple index is the nature of the distribution of scores obtained by the individuals of the group. If the test is too difficult most individuals will tend to get low scores. If it is too easy most individuals will tend to get high scores. The difficulty of a test should be so gauged that the scores show a reasonably wide distribution. In Fig. 7-1 are presented distributions of scores of a group of men and a group of women on a 40-item test designed to measure mechanical aptitude. The scores for the women show a wide distribution indicating that in terms of difficulty, at least, the test was satisfactory for them. On the other hand, the scores for men cluster around the upper limit of the test, indicating that it was too easy for them. A test that is too easy fails to discriminate among individuals of high ability. A test that is too difficult fails to discriminate among individuals of low ability.

It has already been pointed out that, in general, speed tests are not suitable for older persons. To some extent this also holds true for individuals with language handicaps. It is important to fit the verbal factors of the test to the verbal facility of the group. For workers with severe language limitations a non-verbal test should be used. In some tests the language is not cast in terms of the cultural and educational levels

of the workers under consideration. Less educated workers will not react favorably to tests involving difficult words or complex expressions. It is good practice to fit the test to the group in as many ways as can be foreseen.

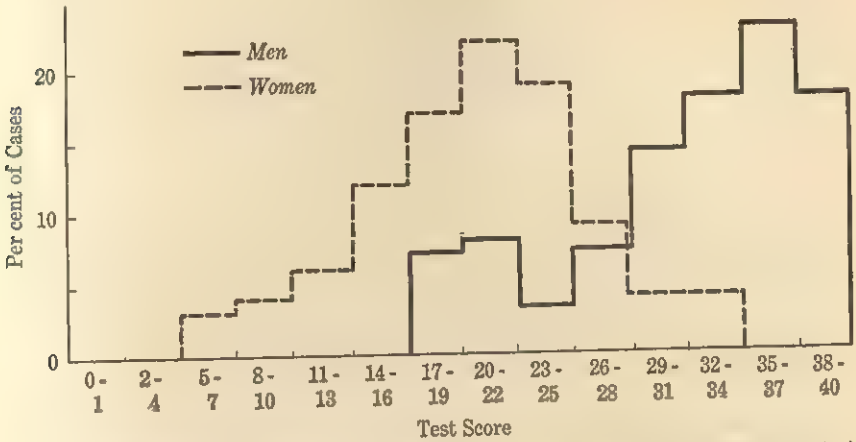


FIG. 7-1. Distribution of scores of a group of men and a group of women on a test of mechanical aptitude.

Norms. The value of a test is largely dependent upon the norms that are available for it. Norms refer to distributions of scores on a test that can be used as a basis for interpreting and evaluating an individual's performance on the test. Usually the test constructor collects scores from representative groups for which the test is recommended, and furnishes these scores to all users of the test.

Several factors determine the usefulness of norms. First of all, other things being equal, the larger the number of workers tested in any group the more dependable the norms will be. Until several hundred scores have been obtained, the norms can usually be considered unstable.

A second factor concerns the amount of available information about the level of skill characterizing the workers whose scores are used in the preparation of the norms. Even when large numbers of scores are available the description of the norming population may be inadequate. Without adequate information about the level of skill of this population the user of the test is unable to assay the test accurately for his specific purpose. It is not sufficient, for example, to say that for a certain test the average score of 200 carpenters is a given value. The classification of carpenter might contain individuals varying widely in ability from apprentices and helpers, who can merely drive nails and saw boards, to master carpenters, who with a given set of blueprints can proceed to

construct a complete building. Unless the degree of skill and knowledge of the group of carpenters is described, the user of the test cannot make an accurate evaluation.

A third factor concerns the kinds and number of different groups for which scores are available. As pointed out several times, the specific use to be made of a test determines the kind of evaluation needed. Unless there is a normed population similar to the group on which the test is to be used, the tester is unable to learn readily the value that the instrument might have for his purpose.

There are two general types of norms: general population norms and special group norms. General population norms are constructed from the scores of workers from a wide variety of jobs. The workers whose scores go to make up these norms are chosen so that they represent a cross section of the total adult employed population of the entire country.

The second type of norms comprises scores of workers from those specific population groups for which the test was designed and for which it has definite usefulness. In Table 7-4 are shown the norms on a test

Table 7-4. Norms on a Test of Eye-Hand Coordination for a Representative Sample of Employed Women and for a Sample of Women Typists

% earning lower scores	Scores for the general employed population	Scores for typists
90	174	186
80	162	178
70	153	171
60	147	163
50	140	157
40	133	150
30	126	144
20	115	134
10	104	123
Number of cases	450	129

of eye-hand coordination for a group of 129 women typists and for a group of 450 women who formed a representative sample of the adult employed population of women. It will be observed that the performance of the typists on this test was superior to that of the general population.

Thus, only 25 per cent of the typists earned scores lower than 140, whereas 50 per cent of the general employed population earned scores lower than this value. This finding is not surprising in view of the fact that this test has some validity for predicting speed and accuracy in typing.

Standardization. The usefulness of a test is determined by the extent to which it is standardized. Any knowledge about a test that is in its favor and has resulted from an empirical application of it improves its standardization. Prescribed and definite methods and procedures for administering a test and for scoring it are part of the standardizing process. Likewise, establishing norms for different groups, as previously discussed, is another important aspect. In so far as the practical situation is concerned, the most significant fact about a test is whether or not it has been adequately validated for the specific purpose for which it is to be used. Because test users have placed so much emphasis on the importance of standardization, a more detailed consideration of this concept is warranted at this point.

Misunderstandings have arisen because of a rather loose interpretation of the term standardization. Three rather widespread notions about standardized tests need discussion. One notion is that a test is a good test because it is standardized; the second is that a standardized test is better than a nonstandardized test; and the third is that a test is standardized when a scoring formula and a time limit are prescribed and a few hundred scores furnished.

Concerning the first notion, it should be said that a test is not a good test just because it is standardized. Rather it is a good test because it measures or predicts the ability or traits for which it is to be used, *i.e.*, it is valid. Of course a test in which the administrative directions, time limits, scoring procedures, etc., are adequately worked out, *i.e.*, standardized, will have a higher chance of proving valid than if these characteristics are not empirically established.

The second notion, that a standardized test is better than a nonstandardized test, may be either true or false depending on what is achieved in the standardization. Suppose there are two tests for selecting bus drivers. The first one has three parts measuring the three abilities of arithmetic computation, social perception, and simple speed of reaction. Each part is timed exactly, the administrative directions are worked out to the last detail, and scores are supplied on 1,000 bus drivers in one transit system who are considered successful on the job. The second test measures two abilities: eye-hand coordination and speed of perception. The administrative directions are rather general, the timing of the parts is not exact, and norms are available on only 200 bus drivers. In addition, however, a validity study was conducted on these 200 drivers. They were

given the test as applicants and then hired. Later such measures as number of accidents, time lost on runs, and number of public complaints were obtained and formed into an index of proficiency. The test scores were then checked against this criterion, the resulting coefficient of correlation being of the order of .50. Some authorities would consider the second test less standardized than the first test. For the purpose of selecting bus drivers, however, the second test is to be preferred to the first.

The third notion involves more directly the meaning of the term standardization. A survey of the standardized tests that might prove useful in business and industry reveals that, with a very few exceptions, standardization refers principally to administrative routine and norming. Validation is not usually part of the procedure, and, in most cases where it has been considered, no great emphasis is placed on it. In terms of the practical usefulness of any test—whether called standardized or nonstandardized—the question is whether the test measures or predicts the ability for which it is intended. Surely, if great care is taken in perfecting the mechanical features of a test, even greater concern should be shown in determining whether the instrument actually does the work for which it is designed. Of all features of standardization, validation would appear to be the most important.

PERSONALITY MEASUREMENT

Success on a job is not solely determined by ability but is also attributable in part to traits of personality, character, and interest. A number of investigations have shown that separation from a job is often due to deficiencies in personality. Consideration should therefore be given to the various methods that have been devised for evaluating these more intangible determinants of job performance.

Definition of Terms. Personality refers to those traits of the individual or those aspects of his behavior that have emotional, social, motivational, or moral connotations such as stability, extroversion, perseverance, and honesty. Numerous authors have attempted to classify the various aspects of personality, and from a theoretical or systematic point of view these classifications are useful. For the problem at hand, however, it is unnecessary to consider such theoretical problems. It will be more fruitful to consider procedures that are employed in developing measures of personality.

General Features of Questionnaires and Inventories. Perhaps the most common way of assessing personality, at least when large numbers of individuals are involved, is by the use of questionnaires or inventories. The terms questionnaire and inventory will be used interchangeably as no sharp distinction in meaning is made between them. A typical inven-

tory consists of a series of questions directly concerned with personality in its behavioral aspects. Following are some sample items:

> Have you ever felt uncomfortable in a social situation?
> Do you often feel just miserable?
> In talking to large groups of people do you get embarrassed?
> Do you object when a person pushes ahead of you in a waiting line?

Depending upon the particular questionnaire or inventory, the answers called for might be simply yes or no, or some qualification in terms of the degree or frequency of occurrence of the behavior might be permitted.

Preparation of a Questionnaire or Inventory. The first step is simply to collect items that seem to have some pertinent relation to the personality trait in question. Various sources of information may be used. Thus Woodworth, in assembling a test for use in the measurement of general emotional adjustment, began by studying case histories of emotionally unstable individuals and recording their reported symptoms.[11] From these symptoms, questions were prepared that could be answered by either yes or no.

The second phase in the preparation of a questionnaire or inventory is validation. An all too common method of validating these personality measures is by proclamation. The constructor of the test states that his test is valid because the items it contains directly pertain to the personality trait in question. Thus a "sociability inventory" is claimed to be valid simply because it contains questions pertaining to behaviors, interests, and feelings which are manifest in social situations. Usually the questionnaire or inventory is based on some "theory" of personality, and the claim is made that a given test is better than most others because the theory on which it is based is a "better" theory. As a science, industrial psychology will be willing to consider such theories as springboards for research, but it cannot accept them as scientific evidence. Various methods of validating personality questionnaires have been employed such as comparing the scores they yield with ratings on personality or observing the differences in scores of groups of individuals whom psychiatrists or clinical psychologists have diagnosed as "normal" and "abnormal" with respect to the trait in question.

Validating an Interest Inventory. Special reference should be made to the validation of one particular type of questionnaire that falls in this general area of personality measurement, viz., the interest inventory. The items in this type of questionnaire usually have to do with likes and dislikes with respect to different occupations, activities, amusements, school subjects, etc. The person answering the inventory may be asked to state whether he likes, dislikes, or is indifferent to each of a series of vocational activities. Or items may be arranged in pairs and the testee asked which of each pair of items he prefers. The items shown in Table 7-5 are typical.

In such an inventory, questions that presumably indicate interest of a given kind are pooled together to give a score measuring that interest.

Table 7-5. Examples of Items in an Interest Inventory

For each of the following, circle L if you like that kind of work or school subject, circle D if you don't like it, and circle I if you are indifferent to it or don't know.

L I D Architect
L I D Barber
L I D Machinist
L I D Typist

L I D Algebra
L I D Biology
L I D History
L I D Mathematics

In each of the following pairs check the one you would rather do.

_____Repair a clock
_____Draw a picture

_____Meet new people
_____Read a book

_____Play baseball
_____Play solitaire

In validating the inventory, workers in a particular occupation are given the questionnaire and their scores in various interests determined and compared with the scores of workers in other occupations. Evidence of validity is obtained when a group of interests is found that appears to be characteristic of successful workers in a given occupation, their scores on these interests being higher than the scores made by workers in other occupations. For example, garage mechanics probably would indicate that they have an interest in the occupation of machinist but no interest in that of actor, that they preferred physical to social sciences in school, and that they would prefer to find out how a radio works than lecture to a group of people. The opposite preferences on these items would probably be checked by a group of salesmen.

Validity as a Function of the Accuracy of Response. The validity of a personality questionnaire or inventory is a direct function of the honesty and accuracy with which the individual answers it. It cannot be expected that a person's performance on a test or inventory will be entirely divorced from the purposes and motives that actuate him to take it. Usually in an ability test a "good" score is a high score. This the individual knows, and since he wants to make the best showing possible he performs as accurately and as rapidly as he can. In the case of per-

sonality and interest inventories, however, a good score is not so defined. Actually on most inventories there is no single "high" score, but the answers are scored in several different ways to obtain scores on several dimensions of personality. It is apparent that an applicant for the job of salesman will not answer "no" to the question "Do you like to meet people?" The individual taking the inventory knows that his performance is going to be used for determining his fitness for the job, and he wants to make a creditable performance. Consequently, he may be inclined to falsify his responses if it appears desirable. Even if he does not deliberately lie, it is likely that his responses to the same inventory would be different on two separate occasions. This might readily occur from different biases aroused by differences in the job requirements when he applies for more than one job.

Sometimes claims are made for particular inventories that falsification on them is impossible. Such claims are not generally borne out by the empirical evidence. Again, when the possibility of falsification is admitted, the claim is often made that it can be detected and measured. Even if such detection were possible, it would take considerable time to measure it for each individual testee. The responses of every person to whom the questionnaire was administered would have to be carefully examined. This examination would have to be done by a highly trained professional person and could not be left to an ordinary personnel interviewer or clerk.

The possibility of falsification does not mean that personality questionnaires or inventories have no place in industrial personnel practice. They are particularly helpful in classification and placement. Suppose that an applicant was assured of being hired and the only remaining problem was to select a particular type of work for him. If he understood that the remuneration and promotional possibilities were the same in all available openings, then it would be clearly to his advantage to give honest responses. The more accurately he answered the items the greater would be his chances of obtaining a correct placement and of achieving success on the job.

The Strong Vocational Interest Blank. Perhaps the best example of a good inventory is the Strong Vocational Interest Blank.[9] It was not designed specifically for use in selection, but rather for use in counseling and guidance. The items in this inventory encompass a wide variety of activities and situations. The norms and validation are based upon large numbers of cases. Furthermore, the establishment of the norms and the validity is a continuing process so the result is an instrument of exceedingly high value when used correctly.

It is perhaps pertinent to note that the Strong inventory, which has an excellent reputation among professional persons, differs in three very

important respects from certain other inventories. First, the scoring key for the inventory is not secret but is available to any professional person. Secondly, the evidence relative to validity is openly published in professional journals and books and is available for all to see. Finally, perfection or even near perfection is not claimed for the inventory. There are other personality and interest questionnaires similar to the Strong inventory in these respects. Like the Strong inventory, they are generally offered as tools for use in guidance and research, and are not offered to industry as the *sine qua non* of personnel selection methods.

Projective Techniques. As an aid in diagnosing maladjustments in personality, clinical psychologists have developed a variety of devices which have been termed projective techniques. The most important characteristic of a projective technique is the great freedom it allows the testee in the content of the responses he may make. The personality inventories described above present specific stimulating situations or problems to the individual and therefore the kinds of responses he can make are determined by the nature of the questions asked. Such a situation is called "structured" since the individual is given a definite frame of reference or structure for his responses. In the projective methods the purpose is to present situations which are very weakly structured. Typical projective situations are ambiguous designs such as ink blots, pictures which are ambiguous in meaning, and incomplete sentences such as "All my life I've . . ." In the ink-blot test the individual is asked to tell what he sees in the designs, in the ambiguous-picture test he is asked to tell the story that each picture illustrates, and in the incomplete-sentence test he is told to complete the sentences as rapidly as he can expressing how he really feels. The logic underlying projective procedures is that since the individual has no frame of reference or structured stimuli given to him by which his responses are predetermined, his responses will therefore be caused by the underlying traits, motives, aspirations, etc., of his personality. In other words, his personality characteristics will be projected or manifest by way of his responses to the unstructured test situation.

Projective techniques have some important limitations. Several reports have been written in specialized journals concerning their value in assessing personnel at the higher levels of management. Claims have been made that they have high validity in this type of situation. The evaluations reported, however, were not based on an objective appraisal of the responses elicited. Rather, a highly trained professional psychologist was required to interpret the responses, and on the basis of his interpretation he made ratings of the individuals' personality characteristics. Since the psychologist ordinarily had much other information concerning the candidates, which undoubtedly was background material for his

judgments, his evaluations seldom were based upon the responses to the projective test alone. Furthermore, it has been demonstrated many times that different psychologists will vary markedly in their interpretations of the same set of responses, indicating a low level of reliability for the meanings assigned to the testee's responses. It can therefore be concluded that, at the present time, it has not been demonstrated that projective techniques are highly valid instruments for appraising personnel in business and industrial situations.

The Forced-choice Technique. One of the most promising techniques for personality measurement in the evaluation of applicants in business and industry involves the use of forced-choice items. The forced-choice technique has been described in connection with rating methods. When used as a test the same general procedures are involved as were described earlier.[6] Pairs, triads, etc., of descriptive adjectives, samples of behavior, etc., are formed so that the alternatives of each group describe traits that appear to be equally desirable. The applicant then chooses from among the alternatives the one which he believes best describes him. Falsification of responses is presumably impossible since all of the alternatives describe equally desirable characteristics, and it is difficult for the applicant to guess the "correct" one.

In validating such a test, counts are made of the responses to each item of workers who turn out to be successful on each job and by workers who turn out to be unsuccessful. Success on the job is evaluated by some measure of worker proficiency. For example, the responses of high producers could be compared with those of low producers, or workers rated high in proficiency with those rated low, or responses of persons who stay a considerable time on the job with those who do not. The examples given in Table 7-6 are from a forced-choice personality test de-

Table 7-6. Examples of Items from a Forced-choice Personality Inventory for the Selection of Supervisory Personnel

Item	% of high-rated supervisors checking item	% of low-rated supervisors checking item
Industrious	30	75
Practical	70	25
Enterprising	40	33
Intelligent	60	67

veloped for the selection of supervisory personnel. In the first item supervisors who are rated high by their superiors describe themselves more often as "practical" than "industrious," whereas the reverse is true for those supervisors who are rated low. Hence "practical" would be taken as the "correct" response. In the second item the responses of good and poor supervisors are almost identical and hence this item would be discarded.

The forced-choice technique of measuring personality, while promising, is so new that not much is known about it. Substantial validity coefficients ranging from .32 to .55 have been found for various groups of Army officers.[6] In certain unpublished investigations where the technique was employed in the selection of supervisory personnel, a validity of a similar order was found. Whether or not an applicant can outguess this type of test remains as yet an unanswered question. It seems likely, however, that the effects of such falsification would be negligible.

TRIAL GROUPS USED IN VALIDATING TESTS

Two General Procedures for Validating Tests. For most tests the initial validation study is made on workers already employed on the job. The reason for this is that measures of job success are forthcoming in a short period of time. No time is lost waiting for the testees to be trained and adjusted on the job as is true when applicants are used. Several limitations of this method will be discussed later.

The second general procedure is to test a group of applicants before they are placed on the job, and then check their test scores against their later proficiency. The ideal setup is to have the applicants placed on the job on the basis of the selective procedures then in use, not utilizing the test scores in any manner in selecting the applicants. In this way there will be a wider range of test scores among the workers, and a more adequate determination can be made of the differentiating power of the test. If the test itself is used to eliminate individuals from the job it is never possible to measure exactly the full extent to which the test predicts job proficiency.

The most important requirement imposed on the trial group of workers is that it be representative of the group on which the test eventually will be used. The specific job for which predictive indices are needed must be kept in mind in selecting this trial group. The testees should be of the same age and sex, have similar job interests and attitudes, and in other respects approximate the general characteristics possessed by those individuals for whom the test is being designed. If the validation results are to hold true for subsequent samples of workers, then the initial trial groups must be representative of these later samples.

Using Old-established Workers in Validation. Old-established workers are not representative of job applicants. As stated above, in order to get the most accurate results, a validation study should be performed on a group of testees representative of those on whom the test eventually will be used. In using a group of employed workers this principle is violated, because workers differ markedly from applicants in their ability to do the job. Workers are not similar to applicants in training, experience, and age and are probably not similar in terms of job interest and attitude. It is apparent, therefore, that the results that would be obtained from such a validation study might differ significantly from those that would be obtained from a group of applicants.

Several kinds of error may occur when established workers are used. The coefficients of correlation, by which the prognostic power of the test is measured, will in most instances be lower in value than if a group of applicants were used. If several tests are being validated, and the problem is to discover a method for combining the tests in order to obtain the greatest effectiveness from the battery, then the results again will differ from those that would be expected from a study of applicants. Any norms established will also be in error. The distribution of scores on the test will not be the same as if the testees were just applying for the job. Any critical scores established must then be considered wholly tentative, and should be corrected in the light of the test distributions obtained later from applicants.

It is not to be concluded that little information is gained from a validation study on established workers. On the contrary, it is almost certain that, if a test stands up under this type of validation, it will have even more predictive power when it is used with a group of applicants. This argument becomes obvious when given a little thought. Any measuring device is more successful in revealing large differences than in revealing small ones when both sets of differences come within its range of application. If a test is shown to be valid on a group of established workers, it means that it is capable of distinguishing varying amounts of ability through a somewhat restricted range; and, since the variations in ability will be greater among applicants, the test will have greater discriminatory power when used with them.

Using Applicant-Workers in Validation. Inasmuch as every test must be validated on workers who are representative of the group of individuals on whom it is to be used, applicants should be the group upon whom the final validation studies are made. To be completely representative of the general group of applicants, the testee applicants should be selected at random. One way of doing this is to test every applicant who presents himself to the employment office during a certain period of time, provided that, for this period of time, no peculiarity of the labor market

occurs to make the sample of testees unrepresentative of the applicant group. If possible, as many as several hundred applicants should comprise the validation group.

The test scores should not be used as part of the selection procedure in determining which applicants should be placed on the job. If it is feasible, all applicants who are tested should be employed, at least temporarily. Instruments for measuring performance on the job should have been prepared and should be set to functioning as soon as possible. Any criterion scores that can be obtained during the training of the applicants should also be used.

During the testing of the applicants, data should be gathered on such factors as age, sex, training, and experience so that studies can be made later of the importance of these variables in determining the scores on both the test and the criterion measures.

PROBLEMS OF ADMINISTRATION

Enlisting the Cooperation of the Testees. The problem of establishing rapport with the testees is an important one. It is a less difficult problem to solve with applicants than with established workers. Applicants are usually strongly motivated to qualify for the job that is open to them. They are then usually willing to follow such procedures as are required to demonstrate their proficiency or aptitude for the job. Despite this willingness, however, they should not be "herded" through the selection procedures in total ignorance of the goals to be accomplished. Rather, they should be given an explanation of what the tests are designed to measure. This information will tend to answer their questions, allay their suspicions and anxieties, and encourage them to cooperate fully.

Greater difficulty is encountered in getting the cooperation of a group of old-established workers for a validation study. These workers often interpret their being used as test subjects as an extra duty, falling outside of their normal obligations to the job they hold. As adults they strongly dislike being exposed to situations in which their abilities are to be evaluated. This is especially true when there is competition present as is the case in a group-test situation. In spite of the fact that they are assured that their test performance will not affect their job standing in any respect, most workers will approach the test situation with a certain amount of trepidation.

Various attitudes on the part of the established-worker testees will be found to affect their test performance. The anxiety incident to being examined has been mentioned. In some cases the attitude will be that of fulfilling only the mechanical requirements of the test. That is, the worker will come to the test and go through the motions of answering

the questions but will not give a performance representative of his ability. On the other hand, other workers will feel that the test results may be used favorably or unfavorably in relation to their job status and they will work doubly hard to make a favorable impression. Some of them may even carry this attitude over to their job and try to improve their pace of work, thinking that in the future they may be scrutinized more closely on the job as a result of their test performance.

About the best procedure to follow for creating adequate rapport with established-worker testees is to give them an accurate account of the purposes of the testing program and to make cooperation in the program a voluntary matter. The worker then can form a correct understanding of just what is wanted of him. One important step is to assure him that the test results will in no way affect his job status. If possible, the test constructor should be given full control over the use of the scores so that he can assure the workers that their supervisors or foremen will not have access to them. A further incentive is to offer to interpret the scores to the workers after the validation is completed. To avoid influencing any criterion data being assembled, the workers should not know their scores until the job proficiency measures have been collected. Cooperation in the testing should be on a voluntary basis, and the workers should be assured that refusal to cooperate will have no untoward effects on their job status. Inasmuch as it is somewhat of an imposition on the worker to take the tests, he should either be allowed to participate on company time or be paid for the time that he contributes in submitting to the tests.

By following the foregoing suggestions it is usually possible to obtain the full cooperation of established workers. Failure to obtain their cooperation during the testing will result in getting an unrepresentative performance from them. This, of course, will reduce the validity of the findings to an unknown and unknowable amount.

Use of Standardized Procedures in Administration. A warning should be given that an investigation of validity should never be attempted until the test is in as final a form as can be achieved. By the time validation testing is begun all technical procedures should have been smoothed out and demonstrated to be in order. Defects in clarity of instructions should have been detected on small trial groups. Time limits, administrative procedures, and other features should have received empirical justification in the same way. It is not desirable to make any changes in procedures after testing has begun, owing to the loss of data such changes inevitably entail.

The administration of the test should conform closely to well-accepted standard procedures. In the validation study the test should be administered under the same conditions as those which will be encountered when

it is actually put to practical use. This means that both the psychological and the physical environments under which the established-worker testees perform must be comparable to those environments to be encountered when applicants are tested. If the test, when adopted as part of an organization's selective procedure, is to be given individually to applicants, then it should be given individually to the workers of the validation group and not administered to them as a group test. If the directions are to be given orally to applicants, then they should be given orally to the worker testees. Unless the directions are to be repeated several times to applicants, they should not be repeated when worker subjects are tested. Only that type and amount of assistance should be given to established-worker testees as eventually will be given to applicants.

The physical environment should include those factors which will be present when the tests are used for practical selective purposes. The tests should not be given to established-worker testees in a quiet room or in a room adjacent to roaring machinery unless such conditions are to prevail when applicants are to be tested. If the tests are to be given to applicants at all hours of the day, then it is not sound procedure to test the workers only at one or two particular hours. It is a well-known fact that there are diurnal variations in performance, and therefore the hour of the day for testing should be the same for the two classes of testees. Two other factors to be controlled are the ventilation and lighting of the testing rooms.

The timing of the tests is a very important feature of administration. Obviously, in speed tests the number of questions that an individual will attempt will be directly related to the length of the testing time. The time must therefore be kept constant for all testees if differences in the numbers of items correctly done are to reflect differences in ability.

It should be expected that the total time of the testing will be greater during the validation study than when the tests are used in the practical selective situation. In the beginning all tests contain a certain amount of "deadwood" that can be removed after it is discovered by means of statistical analyses. Similarly, with a battery of tests there will be some tests that will show sufficiently low prognostic power that they can be discarded without appreciable loss. The testing time required in the validation study will be about twice as long as the time required after the tests have been pruned of their useless parts.

Number of Testees Needed for Validation. A requirement of the validating group of testees, whether established workers or applicants, is that it should be of considerable size. No categorical statement can be made of just how many individuals should be utilized, but certain sug-

gestions can be given that may be helpful in forming the trial group. A group of 20 to 30 individuals is too small for a final validation study. Probably 50 to 60 should be considered a minimum when tests are to be validated singly. If several tests are to be formed into a battery, then 100 individuals is a minimum, and 200 or more will give more reliable results. When the available groups are small, but the labor turnover is high, it is possible to accumulate cases from several applications of the test.

Even when a sizable group is available, reductions in the number can be expected as a result of several causes. When using established workers, some members of the group may be absent from the testing due to illness or other causes. Some reduction may result from erroneous test records where the individuals misunderstood the directions and failed to follow stipulated procedures. A further reduction may result from the fact that criterion measures are incomplete and unusable. Experience gained from many validation studies indicates that the number of individuals completing a study for whom there will be usable test scores and criterion data will be from two-thirds to three-fourths of the number originally available. Knowing this, the test constructor should obtain a group of larger size than that which eventually will be needed to supply data for the desired statistical analyses.

ANALYSIS OF TEST RESULTS

There are several purposes for conducting statistical analyses of the data collected in a validation study, and many different techniques for statistically deriving the quantitative meanings related to these purposes. Only a cursory examination of the major objectives will be given here. The interested reader can find more exhaustive discussions of the problems in standard texts on the principles of psychological measurement.[1]

Test Evaluation. Emphasis has been placed on the need for validating any selective test. The determination of the predictive effectiveness of a test can be made whenever there are sets of test scores and some kind of proficiency measures for the workers under study. The correlational type of analysis, which results in a validity coefficient, has already been discussed. There are many other statistical ways for determining validity, so that some notion of the effectiveness of a test can nearly always be derived regardless of the form in which the criterion data are expressed.

Several other test characteristics can be determined from the data collected in the validation study. It is important to know the reliability of the scores. If a reliability coefficient cannot be derived from the available measures then statistical procedures other than correlation should be applied which will reveal the consistency of the test performances.

The difficulty of the test should also be examined. It is relatively simple to determine the percentage of workers getting various scores on the test. From the shape of the distribution of scores and the amount of variation among the scores, rather accurate estimates can be made concerning the extent to which the test is discriminating between workers at different levels of the ability or aptitude. Further study of the scoring procedures can be made. With criterion data available, variations in the scoring formula can be examined in terms of the final predictive effectiveness. It is then possible to derive statistically the scoring formula which will maximize the validity of the test. Other characteristics such as the timing of the test, the format of the test, the method of recording the testee's responses, the directions for administering the test, etc., can be restudied in the light of the data and experience gained in the validation study.

Evaluation of the Test Items. It is obvious that the value that can be ascribed to the total score of a test is a function of the validity, reliability, difficulty, and other characteristics of the individual items. An evaluation of each of the items of a test can be made from the data collected in the validation study. Responses to each of the items can be checked against the criterion measures in order to determine item validity. Several procedures are available for calculating item reliability and item difficulty. From an examination of the item characteristics it is possible to learn which items have acceptable predictive power, which ones can be modified to increase their predictive effectiveness, and which ones should be discarded as worthless. Such an analysis is effective in increasing the validity of the total test. It usually will also result in a shortening of the test with a consequent reduction in the amount of time required of the testees.

Weighting Tests in a Battery. When several tests are available the problem arises of choosing that combination and number which will give the highest prognostic power. Although the validity of a single test is a direct function of the closeness of the relationship between the test and the criterion, the validity of a battery of tests is a function of both the relationship between each of the separate tests and the criterion and the closeness of relationship between each test and every other test. In general it can be said that the correlation between the tests and the criterion should be as high as possible, and the correlation between pairs of tests as low as possible. In constructing a battery of tests the law of diminishing returns is encountered in making additions to the battery, so that the point is reached when the addition of another test does not pay sufficient dividends in prognostic value for the additional work it entails. Again there are statistical methods available for selecting from among

several tests that particular number and combination which will give the maximum practical validity.[1]

REFERENCES

1. Guilford, J. P.: "Psychometric Methods," McGraw-Hill, 1936.
2. Jenkins, J. G.: Validity for what? *J. Consult. Psychol.*, **10**, 93–98, 1946.
3. Jones, M. H.: The adequacy of employee selection reports, *J. Appl. Psychol.*, **34**, 219–224, 1950.
4. Lorge, I.: The influence of the test upon the nature of mental decline as a function of age, *J. Educ. Psychol.*, **27**, 100–110, 1936.
5. Mosier, C. I.: A critical examination of the concepts of face validity, *Educ. Psychol. Measmt.*, **7**, 191–205, 1947.
6. Rundquist, E. A.: Personality Tests and Prediction, in Fryer, D. H., and E. R. Henry, eds., "Handbook of Applied Psychology," Rinehart, 1950.
7. San Bernardino Air Technical Service Command: "Summary of Statistical Studies Associated with the Development and Interpretation of Tests," 1945.
8. Stead, W. H., and C. L. Shartle: "Occupational Counseling Techniques," American Book, 1940.
9. Strong, E. K.: "Vocational Interests of Men and Women," Stanford University Press, 1943.
10. Welch, J. S., and C. H. Stone: How to Build a Merchandise Knowledge Test, Univ. of Minn. Industrial Relations Center Research and Technical Report 8, 1951.
11. Woodworth, R. S.: "Personal Data Sheet," C. H. Stoelting Co., 1917.

CHAPTER 8

Prediction of Occupational Success by Means of Aptitude Tests

A review of past achievements from the use of aptitude tests should reveal the limits within which they have proved effective, and point up directions in which future research can go in order to increase their effectiveness. Aptitude tests are one of the most promising indices for predicting worker success. Their continued use is assured because of the ever-changing nature of occupations in an industrial society and the continuous fluctuations occurring in the labor market. Some value should accrue from a summary of past findings.

In reviewing the use of aptitude tests it is impracticable to evaluate every test. Although thousands of tests have been devised, many of them have been used but once. In fact, although psychologists have been reporting the effectiveness of their tests for nearly half a century, only on rare occasions has any particular test been validated for a given job more than a few times. Another difficulty is encountered in the meagerness of the reports. Many investigators fail to present adequate descriptions of their tests, of the validation groups, or of their statistical findings, making an accurate appraisal of their work impossible. Further limitations relate to the inadequacy of the indices of job proficiency utilized.

In a review of this kind it is necessary to establish certain basic classificatory schemes in order to discover the common trends or elements among the facts. Often such classifications must be quite arbitrary in nature. One classification used in the present review involved the grouping of aptitude tests according to major types. A second involved the grouping of jobs according to the major occupations. A third required the categorizing of the job proficiency criteria.

A further problem arose in connection with presenting the quantitative data. Fortunately the frequent use of the correlational method justified the restriction of the review to those investigations reporting validity coefficients. In order to reduce the sampling error of these coefficients and

to discover over-all trends in the data, the coefficients from related studies were summarized by means of averages.[7] *

Aptitude tests have been used to predict three different kinds of indices of job success, namely, job proficiency, job training, and labor turnover. Examples of indices of proficiency are production records and ratings given by supervisors. Examples of indices of training are grades in occupational training courses and instructors' ratings. The validity of different types of aptitude tests in the prediction of these two indices of job success will be considered for each of the major types of occupations. The prediction of labor turnover has not been studied so thoroughly, and in addition poses certain special problems. Hence this topic will be dealt with separately.

VARIATION IN VALIDITY COEFFICIENTS OF THE SAME TEST

There is one important finding with respect to the validities of tests that should be noted at the outset. Even when the same type of test is evaluated in terms of predicting the same kind of index of success for the same job, but in different organizations, there will be marked variations in the validity obtained.[6] Examples are shown in Table 8-1. Here are given distributions of the validity coefficients for two tests. Validities are reported from 30 different investigations for finger-dexterity tests when these tests were used in predicting the proficiency of bench workers and assemblers. Similarly, validities are reported from 32 different investigations for intelligence tests when they were employed in predicting the success in training of airplane mechanics. In every instance the validity coefficient is based upon at least fifty cases. It will be noted that for each type of test there is considerable variation in the validity. With assemblers and bench workers, for example, a very high validity coefficient of the order of .85 was found in one investigation, whereas in another investigation a low negative relationship was found. The best estimate of the validity of such tests would be the average of the validity coefficients.

The factors leading to these variations in validity are many. In some instances differences in the kind of workers used in the validation studies would account for the variation. In other instances differences in the reliability of the index of job success or the manner in which the index was determined, would be an important consideration. In many instances the causes of variation in the validity are not discernible.

From the above facts it would appear that even though a test has been found to be useful in predicting success in a given job in one organiza-

* The review given here will be concerned only with research investigations conducted in the United States, and covers the period from 1920 to 1952.

Table 8-1. Variation in Size of Validity Coefficients of Tests Reported for Two Occupations

Validity coefficient	Assemblers and bench workers: finger-dexterity tests vs. job proficiency	Airplane mechanics: intelligence tests vs. success in training
.85 to .89	x	
.80 to .84		
.75 to .79		
.70 to .74		
.65 to .69		
.60 to .64	x	x
.55 to .59	x	x
.50 to .54		xxxxxx
.45 to .49	x	xx
.40 to .44	xxx	xxxxx
.35 to .39	xxxx	xxxxxxxx
.30 to .34	xxxx	xxx
.25 to .29	xxxxxxx	x
.20 to .24	xxxx	x
.15 to .19		xxx
.10 to .14		x
.05 to .09	xx	
.00 to .04	x	
−.01 to −.05	x	
Total	30	32

tion there is no assurance that it would be equally useful in predicting success in that same type of job in another organization. This does not mean, however, that little use can be made of these validity coefficients. Despite the variations, it is possible to learn about mean trends by pooling the validities of a test for comparable jobs and indices of job success. This procedure was used in deriving the mean coefficients presented in the tables to follow.

CLASSIFICATION OF TESTS

As indicated earlier, all manner of tests have been utilized in the measurement of occupational aptitude. Variations among testing devices are so great that the effectiveness of only broad groupings of the tests can be considered here. It is possible to classify the aptitude tests that have been used in the selection of workers into four large categories: intellec-

tual tests, spatial tests, motor tests, and tests of various aspects of character. In the following topics the more commonly used types of tests falling in each of these categories are described.

Intellectual Tests

Intelligence Tests. This type of test was the first that was used in the selection of workers, and more information is available concerning it than for any other type of test. It is still the most widely applied test in employee selection. Many kinds of questions are used in intelligence tests including analogies, reasoning, vocabulary, similarities, opposites, common sense, general information, number extension, and arithmetic. Following are some examples of items found in intelligence tests:

Peripatetic means the same as (1) round, (2) wandering, (3) thorough, (4) vascular, (5) tasty.

A salamander is a kind of (1) vegetable, (2) flower, (3) bird, (4) lizard, (5) stove.

If two men can pack 120 boxes in three days, how many days will it take three men to pack 360 boxes?

What is the next number in this series? 2 4 7 11

Water flows downhill because (1) it is slippery, (2) it is heavier than air, (3) it is subject to gravity, (4) it is pulled by the sun, (5) it is attracted to the sea.

Tribulation is the opposite of (1) sorrow, (2) job, (3) underhanded, (4) open, (5) lazy.

"April showers bring May flowers" has the same meaning as (1) Hitch your wagon to a star, (2) Present pains make us recall past bliss, (3) Into each life some rain must fall, (4) Present hardships make for future ease, (5) Trust in God but keep your powder dry.

Immediate Memory Tests. In tests of this kind the individual is presented with some material for a short period of time and he then attempts to reproduce it. In a typical test a list of materials such as five to fifteen place numbers is given on one side of a page, and the individual is required to note each item, turn the page and write it in an appropriate space on the opposite side. Following is an example:

Front Side	Back Side
1. 396829	1. ________
2. 4839093	2. ________
3. 20768395	3. ________
etc.	etc.

Substitution. In these tests the individual is given a code which he must utilize in making substitutions. He can refer to the code as frequently as he desires. The sooner the individual learns the code, of course, the more substitutions he can achieve. Below is an illustration:

1 2 3 4 5 6 7 8 9
F C K A E Y B H D

From the above code make the appropriate substitutions in the following problems.

E H F Y 5 8
— —

D A C B _ _ _

Arithmetic. Since many occupations require arithmetic computations of one kind or another, this type of test has had wide application. Frequently, the problems are presented in the context of the job. Thus an arithmetic test for cashiers might be made up in the form of a series of transactions involving the making of change.

Number Comparison. These tests consist of a series of pairs of numbers, both members of each pair containing the same number of digits. In about half of the pairs the two numbers are identical. In the remainder one of the digits is different. The task is to indicate which pairs of numbers are the same and which are different, as putting the sign + between all numbers that are the same and the sign o between all numbers that are different. Below are some sample comparisons.

57683 _+_ 57683

703821 _o_ 703921

3159274 _+_ 3159274

Name Comparison. These tests involve the same kind of perceptual judgment as the number comparison test but utilize pairs of names rather than numbers, as illustrated in the three following items:

Littlewell, Forester, & Co. _+_ Littlewell, Forester, & Co.

John F. Grayson _o_ John F. Gragson

Lieutenant George Johnson _o_ Lieutenant George Johnsen

Cancellation. Cancellation tests consist of lines of letters or numbers arranged in random order. The task is to cancel out all of the letters or numbers of a given kind, as in the following example:

Cross out the letter k whenever you see it, like this—~~k~~

d m k w k s i t q k b s b t z b t l k p x f

Spatial Tests

Tracing. Tests of this type are intended to measure speed and accuracy in tracing a path. Ordinarily the test is made up in pencil and paper form. The test may consist of irregular paths printed on paper, and the individual follows along a path by drawing a line with a pencil. An example is given in Fig. 8-1.

Location. Tests of this kind require the testee to locate points on a map, to differentiate distances between pairs of points separated by intervening distracting lines, or to make some other similar judgment involving spatial objects. An illustrative item is shown in Fig. 8-1.

Pursuit. In pursuit tests the individual is required by eye alone to follow a winding line from one point to another through a tangle of other lines. Fig. 8-1 gives an example.

Spatial Relations. In these tests the individual is presented with geometric forms concerning which he must make certain spatial judgments. In the item of Fig. 8-1 the several smaller forms in the square on the left when assembled together will have the shape of one of the four forms on the right. In most cases plane figures are used, although sometimes three-dimensional line drawings are presented.

Speed of Perception. These tests are designed to measure the speed with which differences or similarities in figures can be perceived. Sample items of this kind of test are given in Fig. 8-1.

Mechanical Principles. Problems are used for which the solutions require the application of one or another of various mechanical principles. The problems are presented pictorially and are concerned with the mechanical power of levers, the motions transmitted by gears, the directions

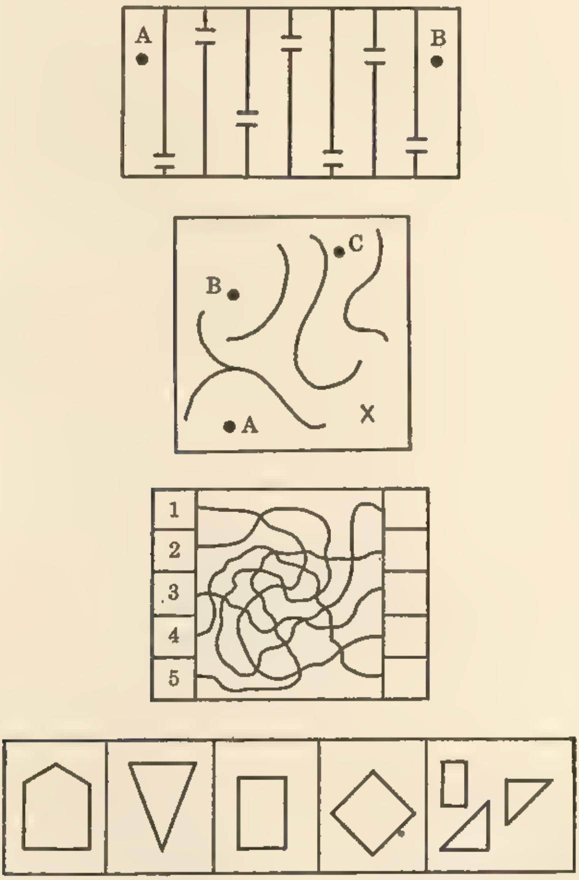


Fig. 8-1. Examples of items in spatial and motor-ability tests.

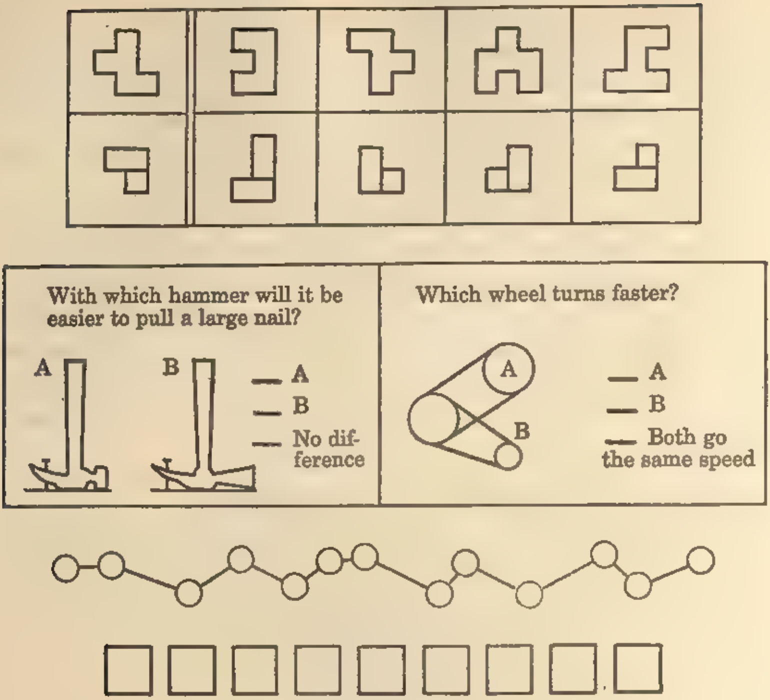


FIG. 8-1 (*Continued*).

and speeds of movements of combinations of pulleys, etc. Examples are given in Fig. 8-1.

Motor Tests

Dotting. These tests emphasize speed and precision of movement. Typically, the individual makes a single dot in each of a series of small squares or circles which are likely to be arranged in irregular order, as shown in Fig. 8-1.

Tapping. Tapping tests are similar to dotting tests except that the emphasis is upon speed alone. As is shown in Fig. 8-1, the circles or squares in which the individual taps with his pencil are sufficiently large so that precision of movement is relatively unimportant.

Finger Dexterity. In most tests of this kind the individual is required to pick up small pegs and insert them into holes. The operation usually takes about 10 min. In some instances, however, he may make a simple assembly such as putting washers on rivet heads and inserting the assemblies into holes, or screwing nuts on to small bolts.

Hand Dexterity. Although these tests involve finger dexterity to some extent, the intent is to sample a somewhat larger set of motions involving the hand and wrist. An example would be a test wherein the individual

picks up a block in each hand, turns them over, and replaces them in the location in which they were originally found.

Arm Dexterity. The intent of these tests is to measure gross motor dexterity involving movements of the arm. They generally involve picking up blocks, carrying them, and placing them in other positions by means of arm movements.

Reaction Time. These tests measure the quickness of response. In a typical test the individual may manually depress a key in response to some stimulus that is presented, such as a light. In the investigations considered in this chapter, reaction-time tests were employed only with operators of motor vehicles. These reaction-time tests usually include a measure of the time required to move the foot from one point to another in response to a visual stimulus. Such a response would be required in the braking of an automobile at the sight of a traffic hazard.

Complex Reaction. Like the simple reaction-time tests, these tests are frequently used with motor-vehicle operators. In complex reactions several stimuli are presented serially to the individual and he must make differential responses to them. Sometimes the score is in terms of the time required to make the responses, but more frequently it is taken as the number of errors made in responding.

Character

Personality. All personality inventories are included in this category regardless of the traits presumably being measured. This is necessitated by the fact that many personality inventories are not labeled in terms of any given trait. Furthermore, in many instances an inventory used in the selection of workers may provide scales for several different traits.

Interest. A number of questionnaires are devised for measuring preferences for various kinds of activities, occupations, etc. Again, in many instances it is impossible to discern the specific interests being measured. All such measures of interests are included in this category.

SUPERVISORY OCCUPATIONS

Employment tests have played a very small part in the selection of persons for managerial and supervisory jobs. This is probably because the proficiency and personal characteristics of individuals considered for these jobs are known either through observation of their behavior in other jobs in the organization or through acquaintance with them in other organizations. Since the work in managerial and supervisory jobs involves considerable responsibility, it is a wise procedure to select people who have been under observation for long periods of time.

One primary difficulty in evaluating tests for selecting persons for supervisory jobs is the lack of adequate indices of success. It has been

difficult to develop for supervisors measures of job proficiency which are both meaningful and predictable. As a consequence, there have been a number of investigations reporting negative results. This does not mean that the success of supervisory workers cannot be predicted by tests. Rather, it means that more experimental research work needs to be done in this area. Since there have been few studies of the effectiveness of tests in the prediction of trainability criteria, results will be given only for the criterion of job proficiency.

Managerial Personnel. Under the term *managerial* all individuals are included who exercise supervision and direction over others with the exception of foremen. Included in this group are store managers, department heads, general managers, etc.

Information concerning the effectiveness of tests in the selection of managerial personnel is extremely limited. As is shown in Table 8-2,

Table 8-2. Average Validity Coefficients of Various Aptitude Tests in the Prediction of Proficiency in Supervisory Jobs

Type of test	Managerial personnel	Foremen
Intellectual:		
Intelligence	.37	.27
Arithmetic		.19
Cancellation	.32	
Spatial:		
Spatial relations		.21
Mechanical principles		.22
Character:		
Personality	.18	.15

data are available for only three types of tests, namely, intelligence tests, cancellation tests, and personality inventories. Surprisingly enough, personality inventories have been ineffective. Of course, it should be obvious that personality characteristics are vital to success in supervisory work. New types of personality measures are needed if valid predictions are to be obtained. On the other hand, both intelligence and arithmetic tests appear quite promising.

Foremen. As is shown in Table 8-2, no particular type of test is outstanding for this group. Again, the personality inventories utilized are found to be of little value. Tests of intellectual and spatial abilities are about equally good and give fair predictions of success.

CLERICAL OCCUPATIONS

Since a large proportion of the employed population is engaged in office work of some kind, tests which may be useful in the selection of employees for clerical jobs are of considerable importance. The nature of the work involved in the clerical occupation varies markedly from job to job, ranging from the relatively unskilled tasks of a messenger boy to the highly responsible work of a chief clerk. Preparation for the latter job may require many years of experience and training and may involve duties of a supervisory nature. For purposes of analysis this occupation is fractionated into three major groups, namely, general clerical workers, recording clerks, and computing clerks. The group of general clerks includes those workers engaged in general office tasks, such as filing, coding, equipment and material checking, classifying, etc. The recording clerks include typists, stenographers, card-punch-machine operators, etc. In the computing group are such jobs are bookkeeper, comptometer-machine operator, and statistical clerk.

Table 8-3. Average Validity Coefficients of Various Aptitude Tests for Jobs in the Clerical Occupations

Type of test	General clerks		Recording clerks		Computing clerks		Average: all clerks	
	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency
Intellectual:								
Intelligence	.41	.38	.40	.25	.23	.16	.35	.26
Immediate memory		.29	.32	.36		.26		.30
Substitution	.21	.31	.24	.24	.34	.27	.26	.27
Arithmetic	.43	.38	.48	.06	.35	.27	.42	.24
Number comparison	.42	.24	.29	.29	.33	.33	.35	.29
Name comparison	.40	.27	.36	.34	.19	.34	.32	.32
Cancellation		.22	.58	.19		.24		.22
Spatial:								
Tracing		−.09	.17	.11		.42		.15
Location		−.03	.24	.11		.28		.12
Pursuit		−.17	.21	.12		.35		.10
Spatial relations	.27	.04	.28	−.02		.30	.28	.11
Speed of perception		.38	.42					
Mechanical principles	.24		.24				.24	
Motor:								
Tapping		.10	.23	.15	.16	.14	.20	.13
Dotting		.11	.13	.15		.03		.10
Finger dexterity		.29	.09	.25		−.11		.14
Hand dexterity		.16				−.05		.06
Arm dexterity		.14						
Character:								
Personality		.29		.19		.13		.20

Effectiveness of Tests in the Prediction of Trainability. As is shown in Table 8-3, intellectual tests are superior to other types of tests as indices of the ability to learn clerical jobs. Tests of spatial and motor abilities, however, are not without some value. It is of interest to note that arithmetic tests give the best indices of trainability for all three clerical occupations. This suggests the importance of quantitative thinking and numerical facility in clerical work. This notion is supported by the fact that number-comparison tests also stand fairly high in validity.

Effectiveness of Tests in the Prediction of Job Proficiency. As far as the prediction of proficiency is concerned, intellectual tests in general show moderately good validities for all three clerical groups. The validity of motor tests tends to be low or negligible. Tests of spatial abilities are of little or no value for general clerks and recording clerks, but are the best type for computing clerks, generally giving substantial results for this group. Personality inventories are fair for general clerks, but have rather low validities for recording and computing clerks.

SALES OCCUPATIONS

For purposes of presentation the jobs classified in the sales occupations are separated into two major groups, salesmen and sales clerks. The group of jobs labeled salesmen refers to those which require a great deal of individual initiative and responsibility such as the jobs of insurance salesmen and wholesale salesmen. The jobs called sales clerks principally involve customer service work and the "over-the-counter" selling involved in department-store work.

Table 8-4. Average Validity Coefficients of Various Aptitude Tests in the Prediction of Proficiency in Sales Occupations

Type of test	Salesmen	Sales clerks
Intellectual:		
Intelligence	.32	−.10
Immediate memory		−.08
Arithmetic		−.12
Number comparison		−.14
Name comparison		−.15
Cancellation		.02
Character:		
Personality	.36	.35
Interest	.32	.34

The scope of investigations concerning the use of aptitude tests for predicting sales work is very limited. The few available studies are summarized in Table 8-4. Intelligence tests and personality and interest inventories have about equal and moderately good validity for salesmen. In the case of sales clerks, tests of intellectual ability almost uniformly have a low negative validity. Personality and interest inventories, on the other hand, give moderately good predictions of proficiency.

PROTECTIVE OCCUPATIONS

Little evidence is available on the validation of aptitude tests in the selection of personnel for the vital work of protecting society, which in-

Table 8-5. Average Validity Coefficients of Various Aptitude Tests for Protective Occupations, Service Occupations, and Vehicle Operators

Type of test	Protective occupations		Service occupations		Vehicle operators	
	Training	Proficiency	Training	Proficiency	Training	Proficiency
Intellectual:						
Intelligence	.46	.26	.50	.07	.18	.14
Immediate memory	.28	.26				
Arithmetic	.30	.08	.59	−.11	.14	.04
Number comparison		.25		.14		
Name comparison		.36		−.17		
Cancellation				−.27		
Spatial:						
Location						.18
Spatial relations	.33	.04	.42		.21	
Speed of perception	.30				.08	
Mechanical principles	.41	.20			.36	.21
Motor:						
Tapping						.32
Dotting						.28
Finger dexterity	.19					
Hand dexterity				−.09		
Arm dexterity				−.01		
Simple reaction time						.27
Complex reaction						.35
Character:						
Personality		.24		.16		
Interest	−.09	−.01				.26

cludes the maintenance of order and protection of life and property. Included here are policemen, firemen, and guards of various kinds. While the number of individuals engaged in this kind of work is relatively small, these workers nevertheless constitute a very important group.

Effectiveness of Tests in the Prediction of Trainability. As will be seen in Table 8-5, intellectual and spatial tests have a substantial validity in indicating which individuals are the most trainable. The only measure of motor ability that has been investigated for protective workers is finger dexterity, and the validity obtained suggests that these tests are of little value. Predicting success in training from interest measures has not proved fruitful.

Effectiveness of Tests in the Prediction of Job Proficiency. Tests of intellectual abilities have fair validities and appear to be the best for the selection of workers in the protective occupations. Spatial tests show little promise. The validity of personality inventories is fair but interest tests apparently have no value.

SERVICE OCCUPATIONS

As with the protective occupations, there have been very few investigations of the use of aptitude tests in the selection of service workers. The personal-service occupations include waiters and waitresses, janitors, and similar service workers.

Effectiveness of Tests in the Prediction of Trainability. As will be seen in Table 8-5, three types of tests have been investigated for these jobs, namely, intelligence, arithmetic, and spatial relations. The validity reported for all of these tests is quite high, indicating that prediction of success in training for personal-service work is effective. One explanation offered to account for these high validity coefficients is the wide range of ability represented among service workers. Persons who are very low in general intellectual ability are often selected for training for personal-service employment. Many are so mentally inferior that they earn very low test scores and also have trouble learning even simple service tasks. These cases produce an exceptionally wide range of ability which would account in part for the high coefficients of correlation. It is a well-known fact that the size of the correlation coefficient is a function of the range of talent for which it is computed.

Effectiveness of Tests in the Prediction of Job Proficiency. Reference to Table 8-5 will show that, with the exception of cancellation, the tests that have been tried out with service workers have low or negligible validity. As was found with sales clerks, a large proportion of the validity coefficients are negative.

VEHICLE OPERATORS

Included among vehicle operators are such workers as streetcar motormen, bus and truck drivers, and chauffeurs. While the nature of the work of these individuals clearly varies from one job to another, they all have in common the operation of moving vehicles within traffic situations.

Effectiveness of Tests in the Prediction of Trainability. As may be seen in Table 8-5 few tests have been studied in relation to the trainability of operators of vehicles. The validity of intellectual tests has been found to be low. Location tests show a little promise. Tests of mechanical principles have moderately good validity and have proved somewhat effective as selection tests.

Effectiveness of Tests in the Prediction of Job Proficiency. In the prediction of job proficiency intellectual tests again have little value. Such information as is available concerning spatial tests indicates that they may have limited value. As might be expected from the nature of the occupation, tests of motor abilities generally have moderately good validity. The validity coefficient for interest inventories shows promise for this type of test.

TRADES AND CRAFTS

The trades and crafts include most of the skilled workers and those occupations that frequently are termed "mechanical." The various jobs that constitute the trades and crafts are listed in Table 8-6. One common characteristic of the jobs in these occupations is that they require considerable training. For many of them formal apprentice-training programs are prescribed, while for others the training is provided by the employer. As with occupations already discussed, considerable variation is found in the validity of different types of tests for one or another of the various jobs. The concern here, however, is with general trends.

Effectiveness of Tests in the Prediction of Trainability. As indicated in the column of Table 8-6 headed "Average," tests both of intellectual (except number- and name-comparison tests) and spatial abilities give fair or moderately good validity for acquiring the skills and knowledges required in the trades and crafts. The best tests are intelligence, arithmetic, spatial relations, speed of perception, and mechanical principles. Motor tests have little or no value for predicting achievement in training.

Effectiveness of Tests in the Prediction of Job Proficiency. The same general pattern of average test validities holds for the prediction of job proficiency as for the prediction of trainability. The major differences are that the validities of intellectual and spatial tests are somewhat

Table 8-6. Average Validity Coefficients of Various Aptitude Tests for the Trades and Crafts

Type of test	Mechanical repairmen		Electrical workers		Structural workers		Processing workers		Complex-machine operators		Machining workers		Average all trades and crafts	
	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency
Intellectual:														
Intelligence	.38	.04	.43	.47	.29	.09	.35	.24	.34	.28	.30	.08	.35	.20
Immediate memory	.30		.31		.13	.13	.31	.15		.30	.12	−.02	.23	.17
Substitution					.31					.26	.27	.21	.29	.24
Arithmetic	.40	.19	.45	.07	.30	.15	.35			.29	.33	.20	.37	.18
Number comparison					−.04	.08	.24			.14	.02		.07	.11
Name comparison					−.01	.08	.14	.20	.20	.22	−.02		.08	.17
Cancellation											.28			
Spatial:														
Tracing	.21		.24	.15		.30	.17	.24	.22	.19	.21	.06	.21	.19
Location	.24		.24	.23	.23	.23	.24	.21	.28	.25	.24	.04	.25	.22
Pursuit	.17		.12	.32				.17		.33		.01	.15	.21
Spatial relations	.34	.19	.33	.33	.28	.31	.35	.16	.36	.30	.33	.11	.33	.23
Speed of perception	.40		.43		.29	.35	.34	.19			.35		.36	.27
Mech. principles	.37	.29	.40		.31		.40			.40	.33	.57	.36	.42
Motor:														
Tapping	−.01			.19	.20	.18	−.01			.19	.05	.08	.06	.16
Dotting	.20				.13	.20	.02			.11	.14	.06	.12	.12
Finger dexterity	.19	.16	.15	.18	.24	.30	.22	.30	.11	.14	.24	.08	.19	.19
Hand dexterity	.17	.12										.29		
Arm dexterity	.08							.32			−.03	.11	.03	.22
Character:														
Personality								.30		.24				.27
Interest												−.13		

lower and the validities of motor-ability tests a little higher. Surprisingly enough, for the two types of jobs for which there are data, personality inventories show some promise.

INDUSTRIAL OCCUPATIONS

In this group are included most of the semiskilled and skilled jobs in industry. The common characteristics of this group are that few of the jobs require any extensive training and most of them involve the manipulation and visual inspection of objects, tools, or equipment. The jobs composing this group and the validities of various tests used for these jobs are given in Table 8-7.

Effectiveness of Tests in the Prediction of Trainability. Since, for most of these jobs, training is of less importance than for the jobs previously considered, it is not surprising to find for training criteria fewer studies of test validity. Intelligence tests appear to have little value, but arithmetic tests have given satisfactory results with at least two groups, assemblers and packers and wrappers. Tests of spatial abilities have fair validity, spatial-relations and speed-of-perception tests showing consistently fair coefficients for the occupations for which data are available. The validities of finger-, hand-, and arm-dexterity tests are moderately high for assemblers.

Effectiveness of Tests in the Prediction of Job Proficiency. With the exception of intelligence and cancellation tests, the validity of tests of intellectual ability tends in general to be low. Among the tests of spatial ability, mechanical principles has moderately high validity. Contrary to expectation for these particular occupations, tests of motor abilities generally show variable but low validities.

CONCLUSIONS CONCERNING THE EFFECTIVENESS OF APTITUDE TESTS IN THE PREDICTION OF TRAINABILITY AND JOB PROFICIENCY

The foregoing review of pertinent findings with respect to the effectiveness of various types of aptitude tests in the prediction of trainability and job proficiency shows that few single tests have high predictive value. Therefore, in the selection and placement of workers, the value derived from the use of single tests will be somewhat limited. By and large, success in training has been better predicted by single tests than has the subsequent level of proficiency attained in job performance.[3]

What the foregoing analysis does not point out is the relative effectiveness of the various tests in measuring different abilities. Differential effectiveness of tests for different jobs makes it profitable to use several

Table 8-7. Average Validity Coefficients of Various Aptitude Tests for Industrial Occupations

Type of test	Machine tenders		Assemblers		Inspectors		Packers and wrappers		Gross manual		Average: all manipulative & observational	
	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency	Training	Proficiency
Intellectual:												
Intelligence		.16	.02	.22	.19	.35	.22	.13	−.03	.26	.10	.22
Immediate memory		.17		.06		.14		.24				.13
Substitution		.19		.12		−.01		.16				.12
Arithmetic		.15	.39	.09		.18	.43	.14			.41	.14
Number comparison		.20		.15		−.02		.13				.12
Name comparison		.17		.10		.17		.20				.16
Cancellation		.25		.36				.24				.28
Spatial:												
Tracing		.16	.16	.18		.20		.12				.17
Location		.11	.29	.19	.19	.18		.16			.24	.16
Pursuit		.15	.28	.15	.09	.09		.16			.19	.14
Spatial relations		.11	.24	.15	.27	.24	.22	.13			.24	.16
Speed of perception			.26	.27	.22						.24	
Mech. principles				.56		.42						.49
Motor:												
Tapping		.12	.16	.14	.10	.06		.14			.13	.12
Dotting		.15	.22	.15		.06		.13				.12
Finger dexterity		.15	.44	.25	.00	.14		.08		.15	.22	.15
Hand dexterity		.23	.50	.14		−.02		.15				.13
Arm dexterity		.15	.54	.24		.00		.24		.43		.21

tests for the same purpose. It would seem reasonable to suppose that combinations or batteries of tests could be formed that would give quite substantial increases in the effectiveness of prediction of occupational success as measured by both trainability and job proficiency. In fact, when such batteries have been developed considerably higher validity coefficients usually have resulted. For almost any of the occupations considered here, and without too much difficulty, it should be possible to develop combinations of tests that would yield validity coefficients of the order of .40 to .50. With more refined analyses, improved indices of job success, and more carefully conducted testing programs, it should be possible to develop selection batteries with even higher validities.

INTELLIGENCE TESTS

Of the many types of tests which have been employed in personnel selection and classification, intelligence tests have been used most frequently and hence deserve special consideration. Intelligence tests have been utilized in the selection and classification of workers for almost every kind of job from the unskilled to the administrative and professional. This extensive use reflects the fact that these tests have frequently been applied uncritically with no attempt being made to understand their true nature. They have been applied and badly misused in situations where they were wholly unsuited. At times the many negative findings have overshadowed the instances where the tests have proved of value. As a consequence, intelligence tests have earned an undeserved reputation of contributing little to the solution of selection and placement problems in industry.

Nature of the Abilities Measured by Intelligence Tests. The major characteristic of intelligence tests is that they are intended primarily to give a very general and over-all picture of an individual's abilities. Like most other tests, intelligence tests measure several different abilities. They differ from other tests primarily in two respects, namely, intelligence tests tend to measure a larger number of different abilities, and the abilities that they do measure tend to be those important in dealing with more abstract and conceptual entities such as those involved in school or academic success. These differences are best shown by the number and types of items which ordinarily are included in an intelligence test. In the same test are found questions on vocabulary, sentence completion, arithmetic, analogies, meaning of proverbs, and reasoning.

If the purpose and nature of intelligence tests are kept in mind, it will be apparent that these tests will have some use in business and industry. But it cannot be expected that they will be of value for all personnel problems. Since there have been many research studies on

intelligence tests, there is a substantial knowledge concerning the kinds of situations in which they will be effective.

In the following paragraphs brief consideration will be given to intelligence tests as indicators of potentiality to learn a job, to produce on a job, to continue on a job, and to be promoted. In addition, the relationship between intelligence test scores and the hierarchy of occupations will be discussed.

Intelligence Tests as Indicators of Potentiality to Learn a Job. For over three decades evidence has been accumulating concerning the effectiveness of intelligence tests in the prediction of school success. For secondary schools, for example, the coefficient of correlation between test scores and grades is of the order of .60. The main consideration for industrial psychology, however, is not the prediction of school success but rather the prediction of success in various trade and technical specialties.

Formal training in the various types of clerical jobs and mechanical trades is generally conducted in institutions of about the secondary school level. While the evidence is not wholly adequate, it appears that the predictive effectiveness of intelligence tests for these applied subjects is somewhat lower than for the more academic subjects. In general, the coefficients of correlation are about .35 for the clerical occupations and for the trades and crafts. In view of the semiacademic nature of the subject matter in training courses in these occupations, the validity of intelligence tests in this respect is not surprising. For the protective and service occupations intelligence tests have been found to have quite substantial validity. The training for the first of these occupations, as is given in police academies, for example, contains much of an academic nature. For the service workers, as has been shown, the exceptionally wide range of ability accounts for the high validity. Finally, as might be expected, intelligence tests are of little value in predicting trainability of vehicle operators or persons engaged in the manipulative and observational occupations.

The above results demonstrate the usefulness of intelligence tests in the prediction of success in formal training curricula. In business and industry, however, most of the training is of a more or less informal sort involving on-the-job training. The extent to which the above findings for formal training can be generalized to cover this type of training is difficult to determine. It would be better, of course, to observe directly the usefulness of intelligence tests in the prediction of success in on-the-job training. As far as the commercial occupations are concerned, there is no satisfactory evidence. For the mechanical trades, there is some evidence from the use of intelligence tests in the selection of apprentices.

The findings in these situations indicate that a validity coefficient of about .35 might be expected.

Intelligence Tests as Indicators of Productivity. When the academic and abstract nature of the content of intelligence tests is considered, it seems somewhat surprising to find that the test scores are related to sheer amount of production on the job. Nevertheless, research findings indicate that intelligence test scores do predict proficiency in production on some jobs. Many of the criteria in the studies of proficiency, from which the data for Tables 8-2 through 8-7 were obtained, were either production indices or largely reflected the worker's achievement on his job. Of course, these criteria were not always units of output since such units are not available for most supervisory and managerial jobs and for some jobs in other occupations. By and large, however, these proficiency criteria can be accepted as measuring production on the job.

Inspection of the validity coefficients of Tables 8-2 through 8-7 will lend support to the following general statements. Intelligence tests have been of particular value in the selection of managerial personnel, general clerks, salesmen, electrical workers, and inspectors. For foremen, recording clerks, protective workers, processing workers, complex-machine operators, assemblers, and gross manual workers the validity of intelligence tests is fairly good. The coefficient of .26 for gross manual workers is not readily explainable on rational grounds. Finally, these tests are of little or no value for selecting computing clerks, sales clerks, service workers, operators of vehicles, mechanical repairmen, structural workers, machining workers, machine tenders, and packers and wrappers.

In interpreting these statements it must be borne in mind that many of the jobs in the trades and crafts and in industry overlap in their specifications and duties. It is hoped that in future research studies in these occupations the investigators will give more definitive descriptions of the particular productive criteria used in validating intelligence tests. It will then be possible to determine more accurately which of the specialized crafts and industrial occupations require the abilities measured by intelligence tests.

Intelligence Tests as Indicators of Ability to Continue on the Job. Ability to continue on the job is one of the most important characteristics of a worker. Every organization suffers some economic loss from employees who leave before they have become fully productive on the job. Adaptation to most jobs involves a period of training during which the individual is nonproductive although being paid wages. In addition, there usually are special costs connected with his training. Therefore, a new worker who completes his training but shortly thereafter leaves the job is a distinct economic loss. Furthermore, his separation from the company is likely to disturb operations through an interruption of work

schedules. Finally, there is the fact that the longer an individual remains on the job the more proficient he becomes and the greater the contribution he can make to his organization. Thus, length of service on the job is an important consideration.

Several investigations have shown that intelligence test scores are related to labor turnover. For many jobs, there is an optimal range of scores within which labor turnover is at a minimum, either higher or lower scores being associated with greater turnover. In some instances, fairly exact minimum and maximum scores can be set. Table 8-8 contains some

Table 8-8. Relation of Intelligence Test Scores to Length of Service for a Group of Cashiers and Inspector-Wrappers

Test score	*Average length of service, days*
90 and above	35
80 to 89	87
70 to 79	96
60 to 69	100
50 to 59	107
40 to 49	142
30 to 39	91
20 to 29	91
10 to 19	3

typical results for a group of inspector-wrappers and cashiers.[10] It will be noted that on the average those workers with intelligence test scores between 40 and 49 remained on the job for the longest time, while those with either higher or lower scores remained on the job for shorter periods of time.

These minimum and maximum scores vary considerably with the nature of the job. In general, they are higher for the more complex jobs

Table 8-9. Relationship between Labor Turnover, Intelligence Test Score, and Difficulty of the Occupation

Grade of work	% turnover of those scoring above 110	% turnover of those scoring below 80
A (hardest)	41	66
B	53	58
C	72	50
D	100	62
E (easiest)	100	37

and lower for the simpler ones. This statement is supported by the data presented in Table 8-9.[2] It will be observed in this table that the proportion of individuals with high scores (110 and above) who leave the job increases as the difficulty of the job decreases. Similarly, the proportion of those individuals with low test scores (80 and below) who leave the job decreases as the work becomes easier.

Intelligence Tests as Indicators of Potentiality for Promotion. Since, in many occupations, workers in lower jobs furnish the pool of personnel who are eligible to be placed in higher jobs, measures of potentiality for promotion are needed. In Table 8-10 figures are given which reveal the

Table 8-10. Relationship between Intelligence Test Scores of Clerical Workers at the Time They Were Hired and Job Level They Achieved Some Time Later

Intelligence test score	Low job	Middle job	High job	Total
	Per cent			
180 and above	4	42	54	100
160 to 179	9	71	20	100
140 to 159	33	59	8	100
139 and below	87	13	0	100

relationship between the intelligence test scores of some three hundred clerical workers obtained at the time they were hired and the level of job they achieved some time later.[3] It is apparent that those individuals who made low scores on the intelligence test at the time they were hired tended to remain on the lower jobs, while those who achieved high scores tended to be promoted to the higher jobs. With clerical and office work, therefore, intelligence tests do give a measure of an individual's potentiality for promotion. Information like that given in Table 8-10 is needed for other types of occupations in which intelligence seems to be an important factor in success.

Conclusions Concerning the Use of Intelligence Tests. It is quite clear from the results cited in the foregoing discussions that the abilities measured by intelligence tests are related to success in learning certain jobs, to the level of productivity achieved on certain jobs, to the length of service attained on certain jobs, and to the level reached on certain jobs through promotion. These generalizations, of course, do not hold for every job. The results on which they are based, however, include such a wide

range of jobs and such a wide variety of intelligence tests that the generalizations are of considerable significance for many problems in the selection and placement of workers.

Intelligence Test Scores as Related to the Hierarchy of Occupations. When the various occupations are to be placed in rank order, it is customary to place the professions at the top of the list, followed in order by the managerial and supervisory jobs, the clerical jobs, the sales jobs, and finally the various manual and mechanical jobs arranged in descending order of amount of skill, viz., skilled, semiskilled and unskilled. It is generally considered that this ranking is based upon the complexity of the tasks involved in the different occupations. It is further thought that the most complex job that a worker can successfully perform is a function of his intelligence. From this it is argued that an individual who performs adequately in a job at a given level can also perform adequately in all jobs which fall at lower levels. A fourth notion is expressed that each job has a minimum intelligence level for success, and that these minima are not the same for jobs at the different levels of the occupational hierarchy. A conclusion is then derived from these several arguments that the relationship between intelligence test score and occupational level is to be explained primarily in terms of the minimum intelligence requirement of the several hierarchical levels. This interpretation should be given a searching examination.

There have been many studies of the relationship between occupational level and intelligence test scores. The same general results have been obtained in all such investigations. Typical findings are presented in Table 8-11, which gives the scores on an intelligence test of soldiers in the Second World War arranged according to their civilian occupation.[9] It will be observed that from the higher to the lower levels of the occupational scale there is a progressive decrease in the average intelligence test score.

The explanation that the minimum intelligence required by each level determines the hierarchical arrangement of occupations should be examined in terms of the variability in intelligence of workers at the different levels. The variation in scores of workers found in any given job is reflected in the values of the first quartile and the third quartile, symbolized as Q_1 and Q_3. Twenty-five per cent of the workers on a job get scores lower than the value of Q_1, and 25 per cent get scores higher than the value of Q_3. In Table 8-11 it will be observed that the variation in scores for workers holding any given job is very large. Thus while the average score of laundry-machine operators is rather low and that of surveyors is fairly high, about 25 per cent of laundry-machine operators

Table 8-11. Scores on an Intelligence Test as Related to Occupation

Scores			Occupation
Average	Q_1	Q_3	
129	121	136	Accountant
124	118	132	Lawyer
122	114	131	Chief clerk
122	114	129	Bookkeeper
120	109	127	Draftsman
119	109	126	Postal clerk
117	107	127	Cashier
117	108	125	General clerk
115	107	125	Salesman
114	102	123	Airplane mechanic
112	103	120	Musician
112	101	123	Toolmaker
111	97	120	Printer
111	97	120	Surveyor
110	100	122	Motion-picture projectionist
110	99	120	Machinist
109	96	118	Policeman
109	95	119	Sales clerk
108	97	118	Locomotive fireman
108	92	119	Cabinet maker
107	95	117	Sheet-metal worker
106	92	117	Shop-maintenance mechanic
105	93	116	Railway brakeman
105	86	115	Auto-body repairman
104	88	119	Structural-steel worker
103	87	114	Plumber
102	89	114	Automotive mechanic
102	88	114	Bricklayer
101	86	112	Pipe fitter
100	87	113	Chauffeur
100	85	112	Wood-working-machine operator
99	83	113	General painter
99	83	113	Baker
98	83	111	Heavy-truck driver
97	79	107	Construction-machine operator
96	79	111	Cook
95	80	109	Light-truck driver
93	82	111	Laundry-machine operator
93	79	109	Barber
93	77	109	Shoe repairman
90	74	104	Railway section hand
87	74	104	Teamster
87	75	103	Miner
86	70	103	Farm worker
85	70	100	Lumberjack

earn scores better than the average of surveyors and 25 per cent of surveyors earn scores lower than the average of laundry-machine operators.

Studies of successful workers scoring low on intelligence tests support the notion that the minimum intellectual levels below which various jobs cannot be performed successfully at all are quite low. Examples of minimum required scores are shown in Table 8-12.[1] It is apparent that persons

Table 8-12. Minimum Mental Ages for Several Jobs

Mental age, years	Boys	Girls
5	Dishwasher	Sewer (simple patterns) Vegetable parer
6	Mixer of cement Freight handler	Mangle operator Crocheter (open mesh)
7	Painter (rough work) Shoe repairer (simple tasks)	Cross stitcher Hand-iron operator
8	Haircutting and shaving Gardener	Scarf-loom operator Dress maker (not including pattern work)
9	Foot-power printing-press operator Mattress and pillow maker	Fancy-basket maker Cook (simpler dishes)
10	Sign painter Painter (shellacking and varnishing)	Sweater-machine operator Launderer
11	Storekeeper Greenhouse attendant	Librarian's assistant Power sealer in cannery

with even quite low mental ages may perform certain jobs in an adequate manner. For example, an individual with an intelligence test score no greater than that of a six-year-old child can handle freight or operate a mangle in a satisfactory manner. Since the proportion of persons in the general adult employable population falling below these minimum levels is very small, such minimum intelligence requirements would seem to be an unimportant aspect of the over-all manpower problem in business and

industry. It would appear, from the facts given thus far, that minimum intelligence requirements cannot be considered the primary factor in determining the relationship between intelligence test scores and occupational level.

There are two additional sets of facts which throw doubt on the notion of minimum intelligence requirements for jobs. First of all a relationship similar to that between occupational level and intelligence test score has also been found to exist for tests of very simple psychological functions.[5] Table 8-13 gives the average scores on four tests earned by persons in

Table 8-13. Scores on Four Simple Tests as Related to Occupational Level

Occupational level	Substitution	Number finding	Dotting	Tapping
Professional	72	67	60	69
Managerial	47	44	53	52
Clerical	64	62	56	70
Sales	62	58	50	57
Skilled	46	46	49	54
Semiskilled	47	46	41	43
Unskilled	29	36	30	37

the various major occupational groups. Two of these tests, substitution and cancellation, measure quite simple intellectual functions. The other two, dotting and tapping, measure motor functions. It will be observed in the table that for each test workers in the "higher" occupations generally achieve higher scores than do those in the "lower" occupations.

A second disturbing fact is that the relationship between labor turnover and intelligence test score sometimes cited in support of the minimum-intelligence explanation of the job hierarchy is also found with other types of tests. In Fig. 8-2 are shown data for several hundred taxicab drivers.[4] It will be noted that for the several tests used, including tests of both spatial and motor abilities, those individuals earning either very high or very low scores are more likely to leave the job than those earning scores around the average of the group.

In the light of the data now available, it is not possible to set down all of the factors which might contribute to the relationship between intelligence test scores and occupational level, or to describe the relative importance of the factors which are known to condition this relation-

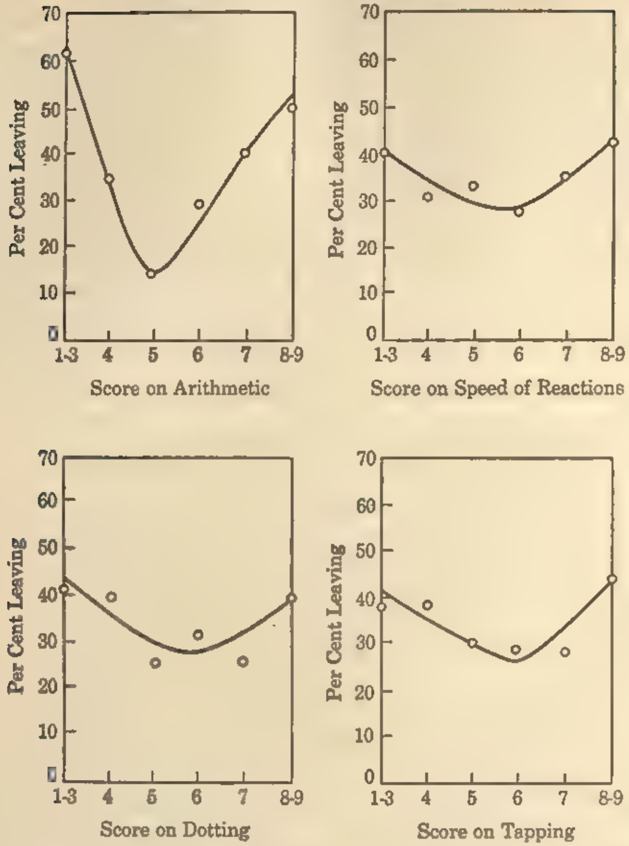


FIG. 8-2. Per cent of workers leaving their jobs in less than three months in relation to scores on four tests.

ship. It does not appear safe to conclude, however, that the relationship is simply attributable to some minimum level of intelligence required for each occupational group.

REFERENCES

1. Beckman, A. S.: Minimum intelligence levels for several occupations, *Personnel J.*, **9**, 309–313, 1930.
2. Bills, M. A.: Relation of mental alertness test scores to positions and permanence in company, *J. Appl. Psychol.*, **7**, 154–156, 1923.
3. Brown, C. W., and E. E. Ghiselli: The relationship between the predictive power of aptitude tests for trainability and job proficiency, *J. Appl. Psychol.*, **36**, 370–372, 1952.
4. Brown, C. W., and E. E. Ghiselli: Prediction of labor turnover from aptitude tests, *J. Appl. Psychol.*, **37**, 9–12, 1953.

5. Ghiselli, E. E., in Kaplan, O. J. (ed.): "Encyclopedia of Vocational Guidance," Philosophical Library, 1948.
6. Ghiselli, E. E.: The validity of commonly employed psychological tests, *Univ. Calif. Pub. Psychol.*, **5**, No. 9, 253–288, 1949.
7. Ghiselli, E. E., and C. W. Brown: Validity of aptitude tests for predicting trainability of workers, *Personnel Psychol.*, **4**, 243–260, 1951.
8. Pond, M., and M. A. Bills: Intelligence and clerical jobs, *Personnel J.*, **12**, 41–56, 1933.
9. Stewart, N.: AGCT scores of Army personnel grouped by occupation, *Occupations*, **26**, 5–41, 1947.
10. Viteles, M. S.: Selection of cashiers and predicting length of service, *J. Personnel Research*, **2**, 467–473, 1924.

CHAPTER 9

Nature and Characteristics of Human Work

Preceding chapters have dealt with problems and procedures connected with the selection, evaluation, and placement of workers on jobs. The next step is to consider problems that arise in connection with the individual's performance of his work. Before discussing the conditions and methods of work it is important to learn something about the nature of the concept of human work. The present chapter will deal with this concept including the psychological and physiological factors associated with work. Factors operating in the physical work space and methods of correctly performing work will be treated in subsequent chapters.

Definition of Work. The common scientific definition of work originated in the physical sciences, where reference is made to changing potential energy into kinetic energy. Work is said to be done when there is movement against a force or when an object is given acceleration. This definition is not readily applied to human work. In many kinds of human work activities, such as solving a problem, little physical energy is involved. To complete the parallel between human activity and physical work, some writers have been led to postulate a psychic energy. During mental activity these writers say that there is an expenditure of psychic energy in addition to such expenditure of physical energy as might occur. However, the concept of psychic energy has not proven fruitful since it is impossible to measure, much less describe.

Further difficulty in carrying over the physical definition of work to human activities arises from the fact that the behavior of human beings is exceedingly complex. Even a job such as that of a fruit packer is not simple. The worker not only makes many hand, arm, and body movements but, in addition, makes a variety of judgments concerning sizes, shapes, and locations, anticipates delays, and plans activities. Human behavior, therefore, cannot be fully described or understood solely in terms of physiological and neurological changes but must be studied at higher levels of a psychological and sociological nature. In addition to being a sensori-neuro-muscular machine, man is an organism capable of conscious thought and reflection, responding in an environment containing

others of his kind. There is no doubt that the higher-level activities, such as thinking, reasoning, and judging, and the conditions in the individual upon which they depend, such as aptitudes, abilities, interests, attitudes, and motives, are all directly related to physiological and neurological processes. However, to be fully understood and explained human behavior must be studied on the level of psychological causes.

If the term work is to be applied to the activities of human beings, the physical definition must be modified. As a beginning, human work can be defined as the exercise of physiological and mental processes in the accomplishment of some end such as a productive operation. This definition is much less exact than the physical definition of work, but since human behavior is so complex a more exact definition is not possible.

Physical and Mental Work. Human work is usually classified as either physical work or mental work. This separation of the physical and the mental is an old and much-debated one. Many arguments have been presented as to whether the mind should or should not be conceived as something different than the body. In studying human behavior the error was readily made of calling the work of the mind "mental" and the work of the body "physical." It is not necessary, however, to introduce the dichotomy of mind and body into the definition of work. All human work involves the physiological machinery of the body, and all work can be differentiated in terms of the extent to which it involves the higher mental processes. Mental and physical work can be separated on a purely descriptive level. Mental work refers to those types of activities in which remembering, thinking, reasoning, judging, etc., are preponderant, and physical work refers to those activities where the speed, coordination, and intensity of muscular responses are the main concern of the individual.

In any specific instance of work this distinction is difficult to make. The work of a machinist, for example, involves movements of the hands and arms but, in addition, it requires the exercise of judgment and memory. It cannot correctly be said that this job primarily involves physical work or mental work. To be sure it is easier to differentiate between the work of an executive and that of a laborer. Even in these cases it obviously would be incorrect to say that the executive job involves only mental work and the laboring job only physical work. In both jobs the individual engages in muscular and motor activities and, at the same time, plans, solves problems, and renders judgments.

The difficulty of distinguishing between mental and physical work has led some investigators to reject this way of classifying human activities. Ryan has pointed out that it is probably more meaningful to speak of sedentary work than mental work, in order to avoid the implication that in mental activity the musculature is not involved.[27] Sedentary work involves problem solving (*e.g.*, calculating, planning), continued sensory

adjustment (*e.g.*, inspecting, radio-code reception), motor skill (*e.g.*, typing, assembling), or light muscular activity (*e.g.*, machine feeding, packing). In sedentary work the control, timing, and direction of activity are more important than in muscular work such as running, carrying, or shoveling, where mechanical force is the paramount consideration. The difference between sedentary and muscular work, however, is one of degree rather than one of kind. This classification of work into sedentary and muscular is not only more realistic but more meaningful than the older distinction of mental and physical work.

SOME CONCOMITANTS OF HUMAN WORK

As the individual works, several more or less characteristic changes occur in the body functions and in behavior. Not all of these changes inevitably appear during work, nor do they occur in the same degree in every individual. Nevertheless, their appearance is sufficiently frequent to warrant consideration.

Physiological Changes. When an individual becomes physically active, a large number of physiological changes occur in the body functions.[22] It is not necessary here to discuss in detail the kinds of physiological changes that occur. The speed of all metabolic processes is increased over that which obtains when the individual is resting. The most readily observable changes are increases in breathing rate, pulse rate, blood pressure, temperature, and oxygen consumption. By use of appropriate techniques the physiologist can also demonstrate changes in the contractibility of the muscles and in the chemical constituents of the blood. Ordinarily the more intense the activity and the longer it lasts, the more extensive all of these changes will be.

Warming Up. In the early part of a working period an individual often shows a gradual increase in his rate of work. To state it in more familiar terms, he starts out cold and warms to his task. This warming-up effect is, in part, physiologically determined. The capacity of the muscles and nerves to perform their functions increases as the activity is continued. Athletes always do some exercising before they enter a contest. No baseball manager risks sending in a relief pitcher unless he has warmed up by throwing the ball for a period of time.

Although warming up has a physiological basis, it also may involve a change in attitude and attention on the part of the worker.[29] It takes him some time to direct his thoughts toward his work and away from other problems with which he was concerned just before he began to work. As he becomes more and more absorbed in his work and focuses his attention upon it, an increase in output results. The phenomenon of warming up is not seen in all curves of work. It is most likely to occur

when the task is discontinuous with frequent short alternations of work and rest.[25]

Decrement in Speed and Quality of Performance. In most situations, as the individual continues to work, both the speed and quality of his performance gradually diminish. Barmack, for example, had subjects perform arithmetical problems for a period of 4 hr.[8] Speed of work was indicated by the number of problems attempted. During the 4-hr. period the number of problems completed decreased 17 per cent and the number of errors doubled. Of course, many factors operate to affect performance during continuous work so there may be instances of performances where no significant changes occur in either speed or quality of output during a continuous application of several hours.

Spurts. The gradual downward course of the curve of work, when performance is represented by the amount achieved, is occasionally interrupted by what appear to be significant increases in the rate of work. Although, in many instances, these spurts cannot be attributed to any known factors, under certain circumstances they occur with marked consistency. With workers paid on a piece-rate basis it is not uncommon to find an end spurt during the last part of the day's work. Again, after rest pauses there are likely to be spurts, and compensatory spurts are found after other types of interruption.[7]

Decrease in Control and Coordination. In some instances, as the individual continues to work, there is a decrease in control and coordination of motions.[1,32] Movements become less precise and accurate. The effects may show up in the form of a lower quality of production. In other instances this lack of control may manifest itself in increased variability in the speed of performance. It is a well-known fact that, during the working day, the individual varies in speed of performance from one time period to another. He works in a series of spurts and slowdowns, alternately increasing and decreasing his speed of work. Late in the working spell, however, these fluctuations are likely to be of greater magnitude. The curve of work then shows an increase in both the frequency and extent of fluctuations as the precision and control of movement decreases with continuous work.

Blocks. If an individual is closely observed during the course of his work, it will be noticed that he stops responding every little while for a very brief period of time.[6] These pauses, called blocks, are very short, being measured in terms of seconds or fractions of a second. The worker apparently is largely unaware of their occurrence. It has been clearly demonstrated that in such mental tasks as arithmetic, reversing perspectives, color naming, substitution, and naming opposites continuous work results in the blocks increasing in both frequency and duration.

After about an hour of work there might be as many as 10 to 12 blocks per minute.

Very little is known about the nature of blocks or the factors producing them. They occur more frequently with homogeneous or repetitive work, and spaced rest pauses reduce both their frequency and duration. The importance of blocks is apparent, particularly with respect to safety. During the period of a block, even though it is quite brief, important events may occur to which the worker is unable to make an immediate response.

Increase in Feelings of Tiredness. A number of investigations have demonstrated the well-known phenomenon that with continuous activity feelings of tiredness increase.[3] The extent of the change in feelings will vary with the nature of the task being performed. With very interesting tasks or with tasks which make small demands upon the individual there may be little or no increase in feelings of tiredness. In addition to the occurrence of such feelings other similar types of mental phenomena may occur. Sometimes there is a sense of strain or of effort and, in very difficult tasks, feelings of conflict or inadequacy may occur.

Change in Attitude Concerning Acceptable Standards of Work. Continuous activity sometimes results in certain other changes that are perhaps best termed changes in work attitudes.[21] Workers who have performed a task for an extended period of time may change their critical attitude toward the quality of their output.[13] Standards of performance are revised. Tolerance for error is raised, and the level of achievement accepted by the worker as necessary and adequate is lowered. The quality of performance accepted as satisfactory late in performance would be considered inadequate by the same worker were he fresh at the task. For example, a pilot starting out on a flight scheduled for 10,000 ft. altitude will make every effort to stay within 100 or 200 ft. of that altitude. After several hours of flying, however, he may consider any altitude between 9,000 and 11,000 ft. as a satisfactory approximation of 10,000 ft. Similarly, the typist who in the morning carefully aligns her typescript and carefully corrects errors, at the end of the day may be satisfied to turn in material poorly aligned and containing errors.

Not much is known about these changes in attitude, and it is quite possible that they do not occur in all types of work. It might be found that changes in attitude toward standards of work are more closely associated with mental than with physical tasks or with boredom and distraction more than with feelings of tiredness.

Recovery in Emergencies. Even after long periods of work, during which there has been little opportunity for rest and the individual is near exhaustion, he frequently can perform surprisingly well during short emergency situations. After days of heavy combat, the soldier can still pull himself together sufficiently for a quick attack. Similarly, the cross-

country truck driver, who has driven his vehicle through the long hours of the night, can safely maneuver it to its destination through heavy morning urban traffic. If the emergency is of sufficient import the individual can muster enough of his resources during the short period of the crisis to perform as satisfactorily as a fresh worker.

THE ROLE OF EFFICIENCY AND FATIGUE IN EVALUATING HUMAN WORK

Traditionally, different methods and conditions of industrial work are evaluated in terms of efficiency and fatigue. A review of the classic writings in the field of human work will reveal numerous references to these two concepts. One method or condition of work is said to be superior to another because it is more efficient or less fatiguing. Very long hours of work are condemned because efficiency drops off and fatigue increases. A new method of work is justified on the basis of claimed increases in efficiency. Rest pauses are instituted on the grounds that they overcome fatigue. High efficiency coupled with low fatigue are the goals sought, and methods or conditions of work that achieve or approximate these goals are designated as being superior to those which fall short of these standards.

When the literature concerned with methods and conditions of work is critically examined, it will be seen that these concepts have not been adequately defined. In many, if not most, discussions of efficiency it will be found that the authors are not discussing efficiency at all. Rather, they are likely to be concerned only with speed of production. Furthermore, fatigue will be found to be quite a nebulous concept. Discussions of fatigue are likely to be characterized by circular reasoning which completely invalidates any conclusions that are drawn. The critical reader will not be content with evaluating human work in terms of these restricted concepts. For example, he will object to the statement that the performance of night workers is wholly adequate because efficiency and fatigue are about the same on the night shift as on the day shift. Certainly he will raise questions concerning the effects upon home life, health, and social adjustment of placing workers on the night shift. The concepts of efficiency and fatigue, therefore, need a closer and more critical examination, and the problem of the meaning and definition of effective work requires a more detailed consideration.

THE CONCEPT OF EFFICIENCY AS APPLIED TO HUMAN WORK

The General Concept of Efficiency. The term efficiency has a very exact definition. It is expressed as the ratio of output to input. Thus a machine that requires 40 units of energy in order to produce 10 units of

work is said to have an efficiency of 25 per cent. A machine that requires the same input, but only produces 6 units, is considered to be less efficient. Applying this argument to human beings, if two workers have the same input but the first produces only half as much as the second, the first is said to be less efficient. Again, if, under two conditions of work, production is the same but, under one condition of work greater human input is required, this condition of work is said to be less efficient.

According to this exact definition there must be knowledge of both input and output before efficiency can be computed. Sometimes the output of two machines is known, but not their input. In such a case their efficiencies cannot be determined, and there is no basis for a comparison on such grounds. Thus, if one machine turns out 20 per cent more completed articles than another, the conclusion cannot be drawn that it is the more efficient since its operation might require 40 or 50 per cent more fuel or power. Similarly, if, under one condition of work individuals produce twice as much as under another, one cannot say that the former condition is more efficient since it might demand three or four times more energy expenditure of the workers. Whether a machine or a man is involved, efficiency cannot be determined for any situations in which there is no information concerning the relative costs required to achieve the production. All that can be said is that there are differences in output.

The Problem of Determining Human Efficiency in Industrial Work. In the industrial situation it is very difficult, if not impossible, to obtain direct measures of human input. One method of measuring the cost of human work is to note changes in the rate of the metabolic processes of the body. To obtain the pertinent measurements it is necessary to attach a rather elaborate and cumbersome apparatus to the worker. This apparatus is likely to interfere with his movements and thereby change the nature of his operations. In addition, the individual to whom such an apparatus is attached needs time to adjust himself to it. If a modification in the working conditions is introduced, the performance will not be equivalent to the actual job performance, and the results may be in error. Furthermore, it seems likely that when a worker knows that his performance is being evaluated his motivation, and hence his performance, is likely to change.[9]

Owing to the difficulties in the working situation of obtaining any measures of physiological changes induced by work, most investigations have been conducted in the laboratory or under semilaboratory conditions. Data obtained in these studies are not usually directly applicable to any particular job in question since the tasks studied are usually not comparable with the job tasks. However, certain general principles about the demands made on the physiological mechanisms by physical work

have been discovered in these studies, and they find pertinent application in many industrial situations.

Many industrial authorities have distrusted the laboratory findings and have been unwilling to learn to what extent changes in human output are bound up with changes in the nature and amount of the input. Most of the time they have considered output alone as a sufficient measure of worker efficiency. When this is done, two important assumptions are made: first, that human input and costs are the same regardless of differences in the methods and conditions of work used, and, second, that the input and costs are the same for all the workers involved. Thus, when a new method of work is found to increase output 20 per cent, it is assumed that the input remains the same, and efficiency is therefore increased by 20 per cent. Numerous investigations have demonstrated the invalidity of this assumption.[9] Certain results, to be cited in the following paragraphs, will show that the physiological costs of work vary markedly under different conditions of work, and that there are wide individual differences in energy consumption on the part of different workers doing the same task.

Examples of Attempts to Measure Human Input and Efficiency by Physiological Costs. A variety of different indices have been used to measure the physiological costs of different methods or conditions of work. Typical indices of body changes are the amount of oxygen consumed and the amount of heat liberated by the individual during work. Sometimes measures of cardiac changes are employed, such as pulse rate or blood pressure. In determining efficiency such physiological indices are compared with the amount of work accomplished by different individuals or with the amount of work produced under different methods or conditions of work. Some typical results obtained from researches in this area are given below.

Sherman studied the energy costs of several common tasks.[28] He estimated the energy used up in terms of the heat expended. The energy required by the resting individual was used as a standard, and the amount expended over and above this was considered the energy cost of the particular task performed. Table 9-1 gives the energy requirements for several activities. It will be seen that typing rapidly required almost double the amount of energy that was expended during resting, and carpentry, metal working, and painting required about three times as much.

Grollman obtained several measures of the physiological costs demanded by the execution of simple arm movements.[15] From his data he computed the relative efficiency of the different movements. The results, given in Table 9-2, show that for two of the indices flexing and extending one arm was slightly more efficient than alternately flexing and extending

Table 9-1. The Relative Energy Expenditure for Several Common Activities

Task	*Relative calories per hour*
Lying still	100
Dressing or undressing	153
Typewriting rapidly	182
Ironing, dishwashing	187
Walking 2.6 miles per hour	260
Carpentry, metal working, painting	312
Walking 3.75 miles per hour	390
Sawing wood	623
Running 5.3 miles per hour	741
Walking upstairs	1430

Table 9-2. The Physiological Costs and Efficiencies of Different Methods of Work

Kind of work	Relative amount of work	Physiological costs			Relative efficiency		
		Oxygen consumption, cc. per min.	Arterio-venous difference, cc. per min.	Cardiac output, cc. per min.	Ratio of work to oxygen consumption	Ratio of work to arterio-venous difference	Ratio of work to cardiac output
Flexing and extending right forearm once per second	100	286	59	4.8	100	100	100
Flexing and extending right forearm twice per second	200	340	65	5.2	167	181	185
Alternately flexing and extending both forearms, each every other second	100	315	73	4.3	91	81	112

both arms. Furthermore, doing this task twice per second was more efficient than doing it once per second. The faster rate of work, while requiring greater input on the part of the individual, did not demand energy expenditure in proportion to the amount of additional work accomplished.

A study of the physiological costs of wheeling a barrow was made by Crowden.[12] This investigator found a wide range of individual differences among men in terms of physiological costs. For four men of about the same skill, the relative amount of oxygen consumed while wheeling a barrow was 100, 161, 171, and 188. Thus, the physiological costs for the fourth individual were nearly twice as great as those for the first. The results of another experiment in which the physiological costs were com-

pared with the rate of wheeling the barrow are given in Table 9-3. It is apparent from these findings that there is an optimal speed of work, with both slower and faster speeds being more demanding.

Table 9-3. The Physiological Costs of Wheeling a Barrow at Different Rates

Rate	Oxygen consumption, cc.	Relative values
Slow walk	1,560	112
Normal brisk walk	1,240	100
Very quick walk	2,040	164
Gentle run	1,960	158

From the foregoing studies, five conclusions can be drawn. First, different types of work require different amounts of energy. Secondly, different methods of accomplishing the same task require different amounts of energy when the energy expenditure is measured by the same index. Thirdly, the relative energy expenditure required for different tasks varies with the different types of indices used. Fourthly, different individuals require different amounts of energy for accomplishing the same task. Finally, the relative amounts of energy required by different individuals performing the same task vary with the method employed in measuring the energy expended.

The fact that different measures of physiological input are not perfectly or even very highly related means that physiological input cannot be thought of as a single variable. It is a complex variable with many aspects, and these aspects are not necessarily highly correlated. Thus, if one index of physiological cost was employed in computing efficiency, the results obtained might be different than if some other index were utilized. The fact that individuals differ in their energy expenditure when doing the same task means that it cannot be assumed that costs are the same for every individual. This is true regardless of the method used in measuring the energy expended. Output, likewise, cannot be accepted as a satisfactory measure of the demands made on the individual in accomplishing a task since it is not highly correlated with input.

Physiological Compared with Psychological Costs. Even if the various physiological indices were found to be highly correlated and to give consistent results, the determination of human efficiency based solely upon physiological costs would still be open to question. Physiological costs are only one aspect of the total demand made upon the individual. Unlike

machines, man is a thinking, reasoning, and feeling being whose activity involves conscious direction. It is therefore apparent that the performance of human work involves psychological costs.

When a worker characterizes a job as "hard work," he may be referring not only to the physical effort he must expend but also to the mental effort, that is, to his subjective feelings of tension, anxiety, or fatigue. In industry it is customary to call a job hard or demanding if it requires a high degree of physical effort, but from the point of view of the individual workman this definition may not be adequate. Psychologically speaking, it is probably harder to inspect containers for flaws and proper labeling as they come down a production line one right after another, making sure that none is missed, than to pack these containers into large cartons and carry the heavy cartons to a truck or a conveyor belt.

The facts indicate that there is little relation between the physiological and psychological costs of human work. In mental or sedentary work the amount of energy expended is so small that it can be detected only by the most sensitive instruments and under the most carefully controlled conditions.[22] Yet light work is frequently described as quite tiring.[31] Even though, by the end of the day, physiological costs to the business executive have been negligible, he can in all honesty report that his endeavors have cost him something personally. Similarly, the activities of the clerk, the instrument assembler, the teacher, and the train dispatcher, while not particularly demanding in terms of physical exertion, exact their toll. Obviously, the ordinary work curve does not measure exactly the costs of the work since the individual may compensate for unfavorable circumstances by exerting greater effort.[30] The interactions between the physiological and the psychological demands of work certainly are highly complex and, to a very large extent, are still unknown.[10]

Psychological Input and Costs. The psychological costs of work, being real, deserve serious attention. They are too readily passed over because they cannot be measured in precise quantitative terms such as calories, degrees, or cubic centimeters. Their existence, however, is quite apparent to the worker. Unfortunately, the psychologist has contributed little to the understanding of these psychological costs. Whereas the physiologist has thoroughly studied body functions in relation to physical costs during work, the psychologist has not only failed to develop instruments for measuring psychological costs, but has not even determined the kinds of costs that might exist.

One kind of psychological cost is called *feelings of effort* or *mental effort*. In some types of activities the individual must intensively concentrate his mental facilities on the problem at hand; in other types the activities run their course with little attention and thus with minimal feelings of effort. Mental effort has been given some recognition in the evaluation

of human work, but thus far little agreement can be reached on an exact meaning for the term. Ryan has sought to clarify the issue by defining effort as the relationship between the actual rate of performance and the capacity of the individual to perform at the given time.[27] While such a definition is useful in that it seeks to specify and clarify a vague concept, it is deficient in that it does not describe the characteristics of the experience of effort itself.

One difficulty with the term feelings of effort is that it is frequently confused with the term fatigue. Fatigue refers to the aftereffects of work, whereas feelings of effort refer to the effects of the immediate demands of the work. If the reader doubts the separate existence of feelings of effort, let him try to multiply mentally two three-place numbers. He will feel little or no aftereffects, that is, fatigue, but unless he is a lightning calculator he will know that he has done some work and that some mental effort was required.

Another psychological cost to the individual in performing his job is emotional in nature. Decisions that involve the welfare of other people cannot be made with the objective abandon that usually characterizes judgments about the operation of machines or the quality of manufactured products. For example, the supervisor who is called upon to lay off men must pay an emotional price for the decision. The executive is faced with severe mental conflict when he must decide how best to adjust his organization to an economic recession. Reducing personnel will probably solve the problem but, in turn, it may remove the immediate means of support of a group of men who have proved loyal employees over the years and who may be the sole support of many other individuals. Such an executive decision is as costly in its way as is the process of solving a difficult problem in the design of equipment. In the first case the emotional cost may far outweigh the feelings of effort engendered in the second. Again, the costs to the individual of such emotionally involved work have not been adequately described, and hence no quantitative evaluations can be made. The person who pays in this fashion, however, certainly is aware of the psychological costs involved.

Mental and emotional costs are as important, if not more important, as physiological costs in the calculation of human input. Under present conditions in this country it is doubtful whether there are many occupations where the work involves such a high physiological cost that it is deleterious to the health of the worker. Legislation relative to working conditions, medical inspection and examinations, and union-management agreements almost ensure that the physical demands on the worker and the physiological costs to him are not excessive. This is not true with respect to the mental and emotional demands on the worker and the resulting psychological costs. Such demands and costs have not been

carefully examined and pertinent data simply are lacking. The depreciation in a machine as the result of its operation can be accurately stated in quantitative terms. Crude estimates can be made of the physiological costs of work to the worker. As yet there are no procedures for evaluating the psychological or mental costs. This is a field toward which much study and research should be directed.

FATIGUE

For at least a century, fatigue has been the subject of many experimental investigations. It might be expected that the relevant problems would have been isolated and defined, and the basic concepts clarified and well understood, but, actually, the issues are still confused, and disagreement is prevalent throughout the field. In many instances the findings from scientific investigations do not agree with common experience; in other instances laboratory findings are not applicable to industrial situations. Thus, although many sound experimental results are available, their interpretation has led to numerous inconsistencies and disagreements.

The General Concept of Fatigue. Ordinarily the term fatigue is used to describe a general state of the individual, a state resulting from continuous activity or work.[26] It is conceived as a reduction in the power to perform or to react. Long periods of activity, the lack of opportunity for rest, and strenuous or heavy work are considered as being conducive to the development of fatigue. Being a general state of the individual, fatigue should show its effects in all aspects of behavior, and all manifestations of it should be of the same order of amount. That is, if one index or sign of fatigue shows a change, then every other index or sign should change in the same manner and to the same degree. Fatigue, then, is thought of as a generalized effect resulting from continuous work.

Fatigue is also thought of as being a causal agent. It is said that fatigue reduces the speed and quality of work, and that it is the cause of accidents. Continued fatigue is held to produce detrimental effects upon health. Here fatigue is being used as an explanatory concept. The term may then be used either to describe a particular state of the individual or to explain certain aspects of his behavior.

The Ambiguity of the Term Fatigue. The definition assigned to fatigue in the foregoing discussion is quite general. However, those who have sought to define it more specifically have come out with varied and inconsistent views. As Bartley and Chute point out, such agreement as exists among different definitions pertains to behavioral changes occurring during work rather than to the phenomenon of fatigue itself.[4]

An analysis of the manner in which the term fatigue has been used in

various discussions of the subject reveals three very different conditions of the individual to which the term is applied, viz., the physiological condition of the muscles and nerves resulting from continuous activity, the reduction in the amount and quality of performance resulting from continuous activity, and the feelings of tiredness resulting from continuous activity. These three conditions are used interchangeably as measures of fatigue. Hence the implication is that they are different manifestations of the same phenomenon—fatigue.

Evidence against Fatigue as a General State of the Individual. It is possible that the phenomena commonly classified as effects of fatigue are specific. That is, as a result of continuous activity there is not just one effect but rather many effects which may or may not be related. In such a case it could not be concluded that fatigue is a general state of the individual. The evidence from pertinent investigations supports this conclusion. The two lines of investigation that are appropriate to consider in this connection are studies of the relationships among different indices of fatigue (decrement in output, physiological condition, and feelings of tiredness) and studies of transfer of fatigue. These two topics will be considered in the following paragraphs.

Relationships between Decrement in Output, Physiological Condition, and Feelings of Tiredness Resulting from Continuous Activity. If fatigue were a general state of the individual then there should be an invariable concomitant variation in decrement in output, physiological condition, and feelings of tiredness. Whenever one state occurs as a result of continuous activity the other two should also occur.

When a reduction in output occurs it is not necessarily accompanied either by an accumulation of the chemical by-products of activity in the muscles and nerves or by an increase in feelings of tiredness.[24] Changes in work output have been shown to be a function of such disparate factors as conditions of illumination and degree of motivation. A reduction in output may result from one of these factors rather than from continuous work and, in such cases, would not be accompanied by an accumulation of physiological products of fatigue.

Studies of the relationship between output and feelings of tiredness show little or no relationship between these two conditions. The lack of any high relationship is common experience. The businessman after a "tiring" day at the office can display a high level of activity on the handball court. Evidence is also afforded by laboratory experiments. Poffenberger had subjects insert words into sentences for a 5½-hr. period.[23] At regular intervals the number of items of completed work was noted and the subjects indicated the extent of their feelings of tiredness on a rating scale. The results indicated that, whereas work output fluctu-

ated irregularly showing no general decrement, the subjects reported increasing feelings of tiredness.

Finally, feelings of tiredness may be unrelated to the extent of physiological changes resulting from work. Benedict and Benedict had subjects do mental work for a period of several hours and measured physiological changes in the body.[5] These experimenters found that the physiological changes resulting from the mental work were very slight. On the other hand, the extent of reported feelings resulting from the work were considerable.

Transfer of Fatigue. Further evidence that fatigue is not a general state of the individual comes from studies of transfer of fatigue effects. If fatigue were a general state, it would be expected that when an individual has performed a task for a sufficient period of time to show a reduction in performance, his capacity to perform other tasks would be equally diminished. The evidence indicates that decrement in performance is highly specific. The extent of transfer of fatigue effects from one task to another varies with the nature of the particular tasks being studied.[8] The greater the similarity between the tasks the greater is the transfer of fatigue effects from the one to the other. Moreover, the transfer effects are seldom complete, so that even with two quite similar activities a marked decrement in one is unlikely to be associated with a similar marked decrement in the other.[27]

Fatigue Tests. Since fatigue is clearly not a general state of the individual it follows that no single index or combination of indices of decrement in performance can be used as a test of the extent of such a general state. Nevertheless attempts are made to compare the general fatiguing effects of different methods and conditions of work by using the extent of decrement in one or more tasks.[33]

For example, in one study it was proposed that the fatigue of truck drivers could be measured in terms of their performance on a series of fairly simple tasks such as reaction time, accuracy of movement, and strength of grip.[18] The argument was advanced that by measuring the decrement in performance on these simple tasks after the men had worked short days, days of moderate length, and very long days, the extent of fatigue from driving could be measured. It was further argued that on the basis of the results of the tests the optimal length of the working day could be determined. Actually on only a few of the tests tried out were the scores adversely affected by lengthening the working day. In other words, the fatigue engendered by driving transferred only to a few other performances.

The use of such procedures to measure fatigue involves an error in circular reasoning. Johnson long ago pointed out the fallacy.[17] After studying many test tasks in which fatigue presumably is to be found, the

investigator then discards all tasks not affected and claims that those tasks deleteriously affected can be used as indices of the amount of fatigue. These tasks are then set up as measures of the fatigue resulting from continuous work. Despite this fallacy in the logic of establishing fatigue tests, investigators still continue their use.

Fatigue an Inadequate Explanatory Concept. A better understanding of the use of the term fatigue will be obtained from tracing the logic underlying its development. The primary fact from which the concept originated is that the worker's performance suffers deterioration with continuous work. This deterioration was noted in the amount of work done, the accuracy of the work, the tendency to have accidents, the feelings of tiredness, the desire to give up or to seek rest, and other similar signs. The physiologist, working first with animals and later with human beings, discovered certain physiological concomitants of continuous work such as increased blood sugar, oxygen consumption, carbon-dioxide production, lactic-acid production, and other similar changes. These changes were also considered as resulting from continuous work.

The word fatigue was coined to stand for the condition present in the individual from which these different changes arose; *i.e.*, it referred to a general state of the organism that caused these changes. This general state is not observable but is known only through the various signs noted above. Fatigue, then, became the cause of the fluctuations in production, the increase in feelings of tiredness, and the changes in the physiological processes. If these diverse phenomena were to be adequately explained as resulting from a single general state, then there would have to be a high correlation among the amounts of their expressions. This has been shown not to be the case. The question arises as to whether there is a single, general state of the organism caused by continuous work that could be described by these various kinds of signs.

It is apparent that, factually, fatigue is a term which merely stands for various effects of continuous work. It follows that, if the effects of the work are diverse and uncorrelated, the term will have several meanings. Because of this ambiguity the term is not useful at the present time, and certain writers have suggested that it be discarded.[20] Others have suggested adopting three meanings according to the nature of the signs observed, *i.e.*, physiological fatigue, industrial fatigue (production), and mental fatigue. When examined closely this is not found to be a satisfactory solution. It was noted above that the various changes resulting from continuous work were not found to be highly correlated. This precludes the use of the term fatigue as standing for a unitary condition within the individual.

A better procedure would be to state exactly the signs or changes being described; *e.g.*, whether drop in quantity of output, increased desire to

stop work, or excessive accumulation of physiological waste products is involved. These are facts, and in the study of continuous work in industry these and similar facts are the sources that will provide the cues for any improvement of methods or conditions of work.

MENTAL ASPECTS OF FATIGUE

Changes in feelings of tiredness resulting from continuous work have generally been given less attention than changes in the amount or quality of work. This does not mean that changes in the worker's feelings are less important, but rather that there is less information concerning them.[26] Certainly there are wide variations in the feelings of tiredness aroused in different workers from the performance of the same job. Furthermore, the extent of a given worker's feelings arising from his job varies on different occasions.

Problems in the Measurement of Feelings of Tiredness. By their very nature, feelings of tiredness are purely private phenomena. Only the individual who is experiencing them can know their extent directly. Others wishing to ascertain the extent of his feelings of tiredness must depend upon his reports. Thus if management were interested in discovering how tiring a particular job was, it would have to depend upon the reports of the workers. Obviously management would regard such reports with some measure of suspicion, since workers might exaggerate the extent of their feelings. Nevertheless, at the present time the only way of measuring these feelings is through an interrogation of the individuals involved.

The usual procedure for measuring feelings of tiredness is to provide some kind of rating scale with steps ranging from "feel very fresh" to "feel exhausted." The worker then reports the extent of his feelings by checking the appropriate point on the scale. This procedure is quite crude and probably very unreliable. Another method is to ask the worker whether he felt more tired, less tired, or equally tired after performing one type of work than he did after performing another. For example, workers who have worked on both the day shift and the night shift might be asked which was more tiring.

Changes in Feelings of Tiredness. A number of laboratory investigations have demonstrated that these feelings increase with continuous activity.[3] Introduction of rest pauses has been found to reduce the extent of such feelings. In industry a similar situation appears to exist. The general trend throughout the working day is toward a gradual increase in tiredness, with maximal subjective feelings reported for the last hour of both the morning and afternoon periods.[14] Cason studied the variations in feelings of tiredness as reported by workers in different occupations.[11]

He found that variations in feelings of tiredness were not the same for all occupational groups. For example, skilled workers reported no increase in tiredness until late in the day, whereas for professional workers there was a marked increased throughout the day. The important finding here is that significant differences in feelings of tiredness are to be expected from different types of work.

Barmack investigated the relationship between feelings of tiredness and the length of the work period.[3] His results indicate that, when individuals are performing continuous mental work, subjective feelings of tiredness probably increase at the same rate regardless of the length of the work period. Although the subjects in Barmack's experiment reported greater feelings of tiredness after longer (4-hr.) than shorter (1-hr.) periods of work, at the end of the first hour those subjects who had only one hour more to work reported feelings of tiredness no different than those who had three hours more to work. This finding was surprising inasmuch as the rate of work of the shorter-interval group was greater than that of the longer-interval group.

Varieties of Feelings of Tiredness. In a most important discussion of the experience of fatigue, Ryan has pointed out that the tired person can and does distinguish several different kinds of feelings.[26] Almost all studies of feelings of tiredness have been concerned with workers' generalized statements about how they feel. Ryan suggests that it would be fruitful to consider characteristic reports of tired persons and to attempt to classify them. This would not only result in a clearer understanding of the nature of tiredness, and hence of fatigue, but in addition would permit a better evaluation of industrial tasks in terms of their demands upon the workers.

In a preliminary analysis, Ryan differentiated seven types of work and endeavored to identify the characteristic reports of feelings of tiredness associated with each. Dromal tasks, such as running, usually result in breathlessness, a pounding in the head, and overheating. When the job is a steady grind, as in continuous barrow wheeling, there are generalized muscular soreness and aches. Certain jobs like continuous typing involve postural restrictions and produce localized stiffness and aches. Strain, trembling, lack of control, and localized loss in strength are frequently reported for repetitive local tasks as in the feeding of machines by hand. With visual inspection and other tasks involving prolonged visual adjustment there is a smarting and watering of the eyes, sleepiness, and headache. Prolonged tasks involving the solving of problems, such as arithmetic calculations, cause sleepiness, fogginess, and postural aches. Finally, situations producing emotive predicaments, such as those involved in the making of important executive decisions, result in general weakness, sleepiness, and residual tensions.

Boredom or Monotony. Another mental state often associated with fatigue is boredom or monotony. Like feelings of tiredness or feelings of effort, it is difficult to define. Monotony is similar to feelings of tiredness in that it is an unpleasant state that the worker seeks to avoid, and similar to feelings of effort in that it is a concomitant of continuous work. Monotony may be defined as satiation for a given activity that the individual is compelled to continue.[2, 34] In essence, monotony arises out of a conflict situation. The individual desires to discontinue an immediate, uninteresting activity, the cessation of which he realizes would endanger the achievement of somewhat more distant goals or rewards. There may be a lack of sense of accomplishment and even a feeling of futility in continuing the work. The feeling tone in boredom is perhaps best described as dull and unpleasant, but such a description is certainly incomplete. It must be admitted that the mental condition of boredom is not easy to describe.

In the industrial situation boredom becomes manifest when the individual finds his work presenting no challenge, and when his interest in the immediate accomplishment of that work is gone. He fully realizes that, unless he continues to produce, he will not be paid. His pay, however, will not be received for a week or perhaps for a month. The monetary rewards for his labor appear to be in the far distant future, and the pleasures to be derived from them are remote, whereas his job is before him, and the unpleasant tasks must be done now. He therefore seeks every dodge to avoid the immediate task without endangering the achievement of his ultimate goal of economic security. He finds himself in a condition of motivational conflict which, though minor, is unhealthful mentally, and its effects on performance will be deleterious if it is of long duration or of frequent occurrence.

Difficulty of Distinguishing between Feelings of Tiredness and Boredom. In many instances it is difficult to distinguish between boredom and feelings of tiredness. Both are the result of work, and both are unpleasant. To produce monotony, work must be carried on for a period of time. Feelings of tiredness result from the earlier phase of an activity and are manifest in the later part of the work spell when monotony is expected to occur. It is apparent then that the criterion of whether the mental state occurs early or late in the work period does not adequately differentiate the two states.

Both states impel the worker to seek a change. The desire to change from boring work may be satisfied by changing to another type of activity, whereas feelings of tiredness move the individual to seek rest. Attempting to differentiate the two states on the basis of whether the change sought is rest or simply a different activity is difficult, since the condition of rest itself is not easy to define. An individual may become

weary as a result of a certain activity and "rest" himself by engaging in some other activity. For example, a person after reading an interesting book for a long period may find that his eyes are tired and his mind somewhat befogged. He takes a rest by walking around the block, chopping wood, or writing a letter. The "rest" in this case is not necessarily either physical or mental inactivity but rather some other type of activity.

The difficulty of distinguishing between boredom and feelings of tiredness has led some investigators to suggest that the two be considered as the same mental state.[19] Certainly many methods that have been found useful in overcoming the one mental state have been found useful in overcoming the other. Nevertheless, on occasion, workers do distinguish between the two. They report that the work on a given job is not particularly tiring but that it is boring; or, that the work is interesting but tiring. In one study of factory workers under conditions of the normal working day, numerous complaints about boredom were found, but there were seldom any reports made about tiredness.[16] It was only when the length of the working day was increased, as in the case of overtime, that complaints about tiredness were made in addition to complaints about boredom.

Boredom as a Function of the Individual. From workers on the same job, reports about their feelings will range from those indicating almost unbearable monotony to those revealing intense interest. Wyatt and Langdon obtained the reactions of workers in four factories concerning the degree of boredom they experienced in their work.[35] Certain of their

Table 9-4. A Comparison of Extent of Monotony Reported by Workers in Four Different Factories

Degree of monotony	Factory A	Factory B	Factory C	Factory D
	Per cent			
Practically free from boredom	1	3	5	3
Slightly affected by boredom	26	42	39	13
Moderate degree of boredom	43	37	33	37
Suffered severely from boredom	28	16	19	30
Never free from boredom	2	2	4	17
Total	100	100	100	100

findings, given in Table 9-4, show the extent of variation among workers. The work of these investigators also indicates that susceptibility to boredom is not a general characteristic of the individual but is a highly specific factor. In Table 9-5 estimates are given by 10 workers of the degree

Table 9-5. Estimates Made by 10 Workers of the Extent of Boredom They Experienced in Performing Five Different Industrial Tasks

Task	Worker									
	A	B	C	D	E	F	G	H	I	J
Packing (14 lb.)	0.4	0.7	3.5	1.4	1.0	1.1	0.9	3.2	0.8	3.6
Packing (4 lb.)	1.5	1.1	4.0	2.5	0.3	1.9	0.5	2.6	1.5	1.0
Staying boxes	1.9	2.0	1.1	1.6	2.1	5.7	3.2	3.0	1.6	1.5
Making crackers	4.0	4.8	1.8	3.7	3.4	1.0	0.3	1.2	2.0	1.9
Bundling chocolates	4.2	4.6	4.3	4.8	3.2	4.4	4.1	4.5	3.8	4.0

0 to 1 Practically free from boredom
1 to 2 Slightly affected by boredom
2 to 3 Moderate degree of boredom
3 to 4 Suffered severely from boredom
4 to 5 Never free from boredom

of boredom they experienced on five different industrial tasks, every worker having spent one month on each task. It will be observed that there is considerable variation within each individual between different tasks. It will also be seen that for a given task there is generally considerable variation from worker to worker. A task that one person considers very boring will be considered by another as almost free from boredom. It can be concluded that susceptibility to boredom is definitely not a general trait of the individual, and that for any particular task individuals will markedly vary in the extent to which they are bored.

THE MEANING OF EFFECTIVE WORK

In previous discussions the concepts of efficiency and fatigue have been shown to be inadequate. The effectiveness of work of human beings cannot be evaluated in terms of either of these concepts. A further difficulty with their evaluative use—a difficulty of utmost importance in industry today—is that these concepts are far too narrow and restrictive in scope. They assume that the worker is solely a producing machine. By

implication, at least, they not only deny the importance of other relations between the individual and his job, but also disregard many of the qualities of the worker as a human being which should be included in any attempt to evaluate his effectiveness.

Some Criteria for Evaluating the Effectiveness of Human Work. The life of the worker on his job is not independent of, and unaffected by, his life away from work; nor is his life off the job unaffected by his activities on the job. Criteria of the effectiveness of a worker's contributions must consider all aspects of the individual, not only as he functions as a worker on the job, but also as he functions as a member of his family and as a citizen of his community. In evaluating the effectiveness of human work it is necessary to establish the standards on a very wide base. Attention must be given to every criterion by which work can be judged as effective or noneffective. The outline below attempts to list the major and more obvious kinds of factors that should be taken into consideration in evaluating the effectiveness of work. Of course, there are so many aspects to the worker's economic, mental, and social "lives" that the list necessarily must be incomplete.

- Economic considerations
  - Amount of production
  - Quality of production
  - Wastage, errors, damage
  - Lost time
  - Labor costs
  - Compensation for unemployment and disability
- Physiological and physical considerations
  - Physiological and physical demands upon the individual
  - Injuries
  - Short-term illnesses
  - Permanent effects on health
- Psychological considerations
  - Feelings of tiredness
  - Feelings of effort
  - Emotional costs
  - Monotony
  - Satisfaction with work, job pride
  - Personality adjustments
- Social considerations
  - Family adjustment and happiness
  - Level of standard of living
  - Cultural achievements
  - Social status
  - Crime and juvenile delinquency

The Complexity of the Task of Evaluating the Effectiveness of Work. It is apparent that any particular method or condition of work can be evaluated in a variety of different ways. It cannot be assumed that, be-

cause one method or condition is superior to another in certain respects, it is more effective in all other respects. Furthermore, it cannot be assumed that certain kinds of indices are necessarily more or less important than other kinds. In most cases, different indices of effectiveness cannot be reduced to a comparable base in order to establish their relative importance. It cannot be said, for example, that an increase of 10 per cent in the number of complaints concerning boredom or feelings of tiredness is more or less important than a 10 per cent increase in the number of units of work produced. The position cannot be defended that benefits in any one respect will invariably counterbalance costs or demands in another respect.

Another factor requiring attention in evaluating any new work method or procedure is the cumulative effects that it may have on the worker. Much of the time the sights of management are set upon immediate increases in production, and the worker's sights set upon the immediate improvement of his occupational status. Thus both are failing to evaluate achievement over the span of productive years that is available. This shortsightedness limits what can be accomplished by both management and the worker. The conservation of human productive power requires an evaluation of energy expenditure over a term of years of worker productivity, and not simply for a given day, week, or month of work. Workers should utilize their energy not only to accomplish the tasks of the day but to continue an effective productivity on the job over the span of years generally considered as years of useful employment. Furthermore, workers should utilize their energy not solely for achievements on the job but for accomplishing the tasks of all other phases of living considered necessary for a well-rounded and balanced life. On the other hand, the worker is obligated to his employer to the extent of not dissipating his energies in off-the-job activities to the point that his effectiveness on the job is lowered. This obligation also must be extended to cover the years of productive usefulness and not be restricted to day-by-day effects.

REFERENCES

1. Ash, I. E.: Fatigue and its effects upon control, *Arch. Psychol.*, No. 31, 1914.
2. Barmack, J. E.: Boredom and other factors in mental effort, *Arch. Psychol.*, No. 218, 1937.
3. Barmack, J. E.: The length of the work period and the work curve, *J. Exp. Psychol.*, **25**, 109–115, 1939.
4. Bartley, S. H., and E. Chute: "Fatigue and Impairment in Man," McGraw-Hill, 1947.
5. Benedict, F. G., and C. G. Benedict: The energy requirements of intense mental effort, *Proc. Nat. Acad. Sci. U.S.*, **16**, 438–443, 1930.

6. Bills, A. G.: Blocking: a new principle of mental fatigue, *Am. J. Psychol.*, **42**, 230–245, 1931.
7. Bills, A. G.: "The Psychology of Efficiency," Harper, 1943.
8. Bills, A. G., and W. McTeer: Transfer of fatigue and identical elements, *J. Exp. Psychol.*, **15**, 23–36, 1932.
9. Bitterman, M. E.: Lighting and visual efficiency, *Illum. Eng.*, **43**, 906–931, 1948.
10. Brozek, J., in Patty, F. A. (ed.): "Industrial Hygiene and Toxicology," Interscience, 1948.
11. Cason, H.: General curves and conditions of feeling, *J. Appl. Psychol.*, **15**, 126–148, 1931.
12. Crowden, G. P.: The physiological cost of the muscular movements involved in barrow work, *Ind. Fatigue Research Bd.*, No. 50, 1928.
13. Drew, G. C.: Fatigue, Air Ministry, *Flying Personnel Research Comm.*, No. 488, London.
14. Griffith, J. W., W. A. Kerr, T. B. Mayo, and J. R. Topal: Changes in subjective fatigue and readiness for work during the eight-hour shift, *J. Appl. Psychol.*, **34**, 163–166, 1950.
15. Grollman, A.: The effect of mild muscular exercise on the cardiac output, *Am. J. Physiol.*, **96**, 8–15, 1931.
16. Jahoda, M.: Some socio-psychological problems of factory life, *Brit. J. Psychol.*, **31**, 191–206, 1941.
17. Johnson, H. M.: Index-numerology and measures of impairment, *Am. J. Psychol.*, **56**, 551–558, 1943.
18. Jones, B. F., R. H. Flinn, and E. C. Hammond: Fatigue and hours of service of interstate truck drivers, *U.S. Public Health Service, Public Health Bull.*, No. 265, 1941.
19. Mayo, E.: Revery and industrial fatigue, *J. Personnel Research*, **3**, 273–281, 1924.
20. Muscio, B.: Is a fatigue test possible?, *Brit. J. Psychol.*, **12**, 31–46, 1921.
21. Myers, C. S.: Conceptions of fatigue and adaptation, *Psychol. Rev.*, **32**, 1–16, 1925.
22. Page, R. M.: Measuring human cost in industry: a general guide to the literature, *Genet. Psychol. Monograph*, **11**, 321–537, 1932.
23. Poffenberger, A. T.: Effects of continuous work upon output and feelings, *J. Appl. Psychol.*, **12**, 459–467, 1928.
24. Robinson, E. S., in Murchison, C. (ed.): "Handbook of General Experimental Psychology," Clark University Press, 1934.
25. Robinson, E. S., and W. T. Heron: The warming-up effect, *J. Exp. Psychol.*, **7**, 81–97, 1924.
26. Ryan, T. A.: Varieties of fatigue, *Am. J. Psychol.*, **57**, 565–569, 1944.
27. Ryan, T. A.: "Work and Effort," Ronald, 1947.
28. Sherman, H. C.: "Chemistry of Food and Nutrition," Macmillan, 1941.
29. Skaggs, E. B.: A study of the "warming-up" in the case of a task of more complicated perceptual-motor coordination, *J. Appl. Psychol.*, **15**, 499–511, 1931.
30. Smith, M.: Some studies in the laundry industry, *Ind. Fatigue Research Bd.*, No. 22, 1922.
31. Walker, C. R., and R. H. Guest: "The Man on the Production Line," Harvard University Press, 1952.

32. Weinland, J. D.: Variability in performance in the work curve, *Arch. Psychol.*, No. 87, 1927.
33. Welford, A. T., R. A. Brown, and J. E. Gabb: Two experiments on fatigue as affecting skilled performance in civilian air crew, *Brit. J. Psychol.*, **40**, 195–211, 1950.
34. Wyatt, S.: Boredom in industry, *Personnel J.*, **8**, 161–171, 1929.
35. Wyatt, S., and J. H. Langdon: Fatigue and boredom in repetitive work, *Ind. Fatigue Research Bd.*, No. 77, 1937.

CHAPTER 10

Conditions of Work and Productivity

A worker's contribution to his organization in terms of his productivity is affected by the various conditions under which he is required to work. Some conditions in the working environment operate to improve his production while others reduce it. An understanding of these conditions will aid in the development of working environments that will enhance the productive possibilities of the worker.

DEFINITION OF PRODUCTIVITY

Output as a Measure of Productivity. Output has two equally important aspects: amount and quality. In studies of conditions or methods of work these must be considered separately since they may vary independently or even in an inverse manner. Under certain circumstances it has been found that as speed of work is increased quality is reduced. Indeed, just as there are differences between different workers or different conditions of work with respect to quantity of output, so there are similar differences with respect to quality of output. Nevertheless, as will be apparent from the following discussion of pertinent investigations, little attention has been given to quality as a measure of productivity. There appears to be a general assumption that all production is satisfactory in terms of quality, since in most instances the work passes a final inspection. But it is to be noted that standards of what is considered satisfactory quality may vary from organization to organization and even within the same organization from time to time. Furthermore, even though the work passes the immediate, arbitrary standard established, there may be variations in quality above that standard which will reflect differences in worker proficiency.

Lost Time as a Measure of Productivity. In studying the effects of conditions of work upon output, the common procedure is to compare changes in average output per unit of time with changes in conditions of work. In computing the averages the total output for a given period is divided by the man-days or man-hours worked. Such a procedure is satis-

factory if the objective is to study only the speed with which the men work.

Under some circumstances, however, a given condition of work may not have any effect upon speed of work but may affect lost time. This time lost from the job, of course, will be reflected in a reduction in the over-all output. Thus, although the output of the employees who are working may remain practically unchanged in terms of hourly or daily production, total or long-term output might be considerably reduced. For example, installation of an air-conditioning unit may increase the comfort of the individual employee so that he can work faster, but if the sudden change in temperature experienced upon entering and leaving the workroom gives him a cold and forces him to stay away from his job for a day or so, total output will suffer. Furthermore, a succession of these colds may easily reduce the vitality of the individual so that a more serious illness may develop. The new condition of work therefore can be said to have an adverse effect upon output, to say nothing of the discomfort it produces in the employee, or of the economic loss that may be incurred either by him or his employer.

There are, of course, many causes of lost time, illness being only one. Strenuous work for long periods, low worker motivation and morale, maladjustments in the home, and other similar factors will cause workers to stay away from their jobs. The analysis of such causes may be very revealing, but the end result is the same—lowered productivity.

The amount of lost time ordinarily is not equally distributed among a group of workers. Most persons lose little or no time, whereas a few workers lose a great deal. For example, among groups of factory workers Vernon found that about half of the lost time was due to one-sixth of the workers, and 80 per cent of it was due to one-third of them.[87]

DIURNAL CHANGES IN PRODUCTIVITY

In Fig. 10-1 are shown the rates of productivity on two different jobs, one in an 8-hr.-day plant and one in a 10-hr.-day plant.[15] It will be observed that the curves have somewhat similar general configurations. There is a tendency in both the morning and afternoon work spells for speed of production to increase to a maximum and then fall off, with the over-all speed of work during the morning being somewhat greater than that of the afternoon. The observation that some daily work curves have similar general shapes has given rise to the notion that there is a "typical" work curve. Although it is probably true that, if the daily work curves obtained from a large number of different jobs were examined and classified, many of the curves would fit the above description, nevertheless, the largest proportion would in one or more respects differ from this

shape.[32] Indeed, examination of Fig. 10-1 will indicate that only one of the four curves—that for dexterous handwork in the 10-hr. plant—closely fits the description given above. The "typical" daily work curve, therefore, cannot be considered descriptive of the variations in performance that occur on most jobs. It is typical only in the sense that it is similar to more of the obtained curves than is any other shape of curve.

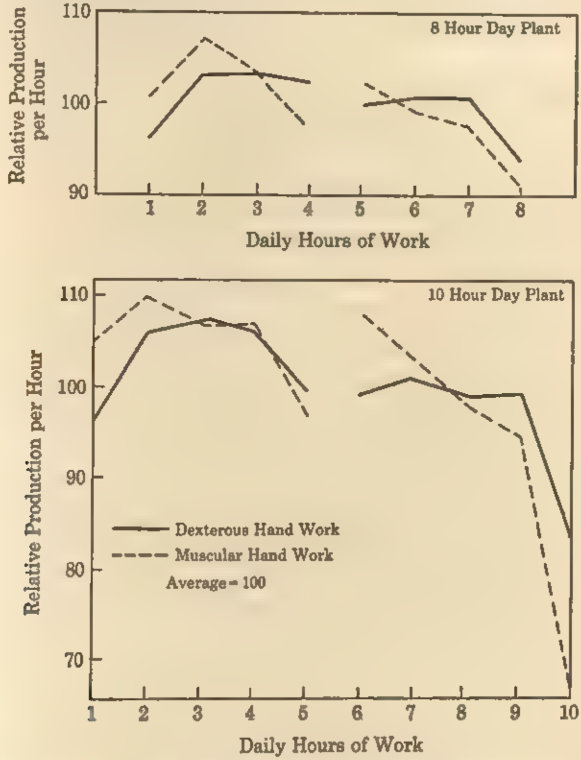


FIG. 10-1. Daily curves of work for dexterous handwork and muscular handwork in an 8-hr.-day plant and in a 10-hr.-day plant.

It is apparent therefore that the daily work curve will vary markedly with the nature of the job. Generally, the more physically strenuous the work and the longer the workday, the greater will be the general reduction in the rate of work from the beginning to the end of the day. Also, with workers on the night shift there is likely to be a greater loss in speed of work from the beginning to the end of the work period than with workers on the day shift. In certain instances, where the rate of work is controlled either voluntarily or because of machine speeds and supply of materials, the daily work curve may appear as a horizontal straight line.

LENGTH OF THE WORK PERIOD

Length of the Work Period and Speed of Production. During national crises, such as war, it has been common practice to lengthen the workday to as much as 14 or even 16 hr. in anticipation of greater production of manufactured goods to satisfy increased demands. The fallacy of the indiscriminate application of such procedures was dramatically revealed in England during the First World War. In a munitions plant, when the

Table 10-1. The Relationship between Hours of Work and Output in Certain Factory Occupations

Length of work period		Relative output	
Hours per week	Relative length	Per hour	Per week
Sizing fuse bodies: heavy work, completely hand			
66.7	100	100	100
60.2	90	120	105
55.5	83	137	119
Turning fuse bodies: heavy work, chiefly hand, partially machine			
74.5	100	100	100
63.5	85	121	100
55.3	75	156	113
Milling a screw head: chiefly machine, partially hand			
71.8	100	100	100
64.6	90	121	102
57.3	80	133	97
Boring top caps: almost wholly machine			
76.8	100	100	100
60.1	78	121	93

working week was reduced from 66 to 48.6 hr., a reduction of 26 per cent, hourly output was increased by 68 per cent and total output for the week by 15 per cent.[36] This finding could be multiplied many times.[18] Typical results obtained in four factory operations are given in Table 10-1.[37] These results illustrate a common finding that heavy work benefits more from a reduction in hours than does light work. In addition it will be seen that the more the work is handwork and the less it involves machines, the greater are the benefits resulting from shortening the length of the work period.

The effects upon production of reducing the length of the working period are closely related to the type of work incentive.[18] Where workers operate under an incentive system in which they are paid according to the number of units of work they turn out, changes in the length of the period of work are less closely related to total production. On the other hand, when workers are paid in terms of the amount of time they work, then production is more closely proportional to the length of the working period.

The pattern of hours worked in terms of their distribution throughout the week is also a factor in determining productivity. A 45-hr. week can be achieved through a pattern of five 9-hr. days or through a pattern of five 8-hr. days plus one 5-hr. day. In general the findings seem to indicate that the number of hours worked per day is a more important consideration than the number of days worked per week.[18] Thus if a 40-hr. week were to be increased to a 48-hr. week, and the original schedule called for five 8-hr. days, it would be better to work an additional 8-hr. day than to increase the hours on each of the five days from 8 hr. to $9\frac{3}{5}$ hr.

Optimal Length of the Work Period. For many industrial operations an 8-hr.-day, 5-day-week schedule appears to be the best.[18] However, no final conclusion can be drawn at the present time as to the optimal length of the workday or the work week. Much more research needs to be done on these problems. The optimal daily or weekly work period varies considerably with the nature of the work, with the type of working conditions, and with the type of worker. In any particular industrial situation this problem will have to be solved by a systematic investigation.

Time Lag in Increase in Production Following Reduction in Work Period. Gains in output resulting from reduced working time will not necessarily be immediately manifest. Some of the results summarized by Vernon are shown in Table 10-2.[37] For operations that are relatively light and are chiefly done by hand, the time to reach a final level of output may be relatively short, as in the case of boring top caps which

Table 10-2. Time Required to Reach Maximal Output after a Reduction in the Length of Hours of Work

Job situation	Change in hours	Number of weeks to reach maximal output	% increase in output per hr.
Women turning fuse bodies	63.5 to 55.3 per wk.	31	29
Women turning fuse bodies	74.5 to 63.5 per wk.	17	21
Men sizing fuse bodies	60.2 to 55.5 per wk.	48	16
Men sizing fuse bodies	66.7 to 60.2 per wk.	9	29
Youths boring top caps	76.8 to 60.1 per wk.	3	21
Open-hearth-furnace men	12 to 8 per day	56	18
Tinplate millmen	8 to 6 per day	8	11

required only 3 weeks to reach an increase in output of 21 per cent. On the other hand, with heavy manual work the time may be considerably longer, as illustrated in the case of open-hearth furnacemen who took over a year to reach an increased output of 18 per cent following a reduction in the length of the working day.

Explanation of Reduction in Output Following Lengthening of Work Period. The decreased output during a longer working period cannot be attributed to a voluntary slowing down on the part of the employees as a protest against the longer working time since the same type of results has been obtained in the psychological laboratory with highly motivated subjects working short periods of time. The worker adjusts his speed of work to the magnitude of the task set for him, slowing up to conserve his energy for a long or arduous task and speeding up when he knows continued expenditure will not be demanded. This adjustment is the same as that made by the track runner. The miler, for example, who started off at the rate of the 100-yard-dash man would likely not complete his race.

Poffenberger reports that, when subjects were instructed to squeeze a dynamometer as hard as possible, they averaged 52 kilograms on the first squeeze when they were informed that they would have to do the task fifteen times, and 68 kilograms when they were informed that they would have to do the task only once.[30] Similarly, Bills and Brown had subjects do arithmetic problems for 1, 2, 5, and 10 min., and found that the shorter the time the subjects expected to work the greater was the output during the first minute of work.[3] These investigators also found that physical effort, as indicated by the pressure the subjects exerted on the pencil

with which they wrote down their answers, was greater when anticipating the shorter than when anticipating the longer periods of time.

Very Short Periods of Work. The conclusion should not be drawn from the findings discussed above that shortening the working time invariably results in increased speed of production. Rather, there appears to be an optimal length of working time, both longer and shorter ones being less effective. Wyatt found that when the working week of cotton weavers was reduced from 5½ to 4 days, the average hourly output fell 6 per cent, and consequently, total weekly output was considerably reduced.[46] Similar results are reported by Miles and Angles for a box-manufacturing plant.[26]

The fact that very short working hours do not bring about greater output is due in part to the so-called "warming-up effect," which manifests itself in terms of a lower speed of operation during the initial period of adapting to the work. The reduced output in the first hr. of work will be more important in a 5-hr. working day than in a 9-hr. one since, in the former, 1 hr. of work constitutes a greater proportion of the working day (20 per cent) than it does in the latter (11 per cent).

In addition, when the worker first comes on the job in the morning, and again after the lunch period, he will lose a certain amount of time adjusting the machines and equipment. Before the end of the working spell he will stop work in order to perform certain other tasks such as cleaning up and putting away tools. This lost time ordinarily is discounted, but it might involve a fair proportion of the total working time. For example, Vernon found that a group of factory workers employed on a 10-hr. day took on the average 14.0 min. to start work in the morning and 12.3 min. to start in the afternoon.[37] They stopped 1.3 min. before quitting time in the morning and 9.3 min. before final quitting time in the afternoon. Thus a total of 37.1 min., or about 6 per cent of the total working time, was not devoted to actual production. With many jobs these nonproductive activities are quite necessary. It seems unlikely therefore that the time they require could be reduced to any great extent. A greater proportion of a shorter working day will be consumed by these activities not reflected in production than in a longer day. This factor, in addition to the warming-up effect, contributes to making output on very short working days inferior to that of somewhat longer ones.

Lost Time and Length of the Work Period. In Table 10-3 data are presented which show the relationship between the length of the work week and the amount of lost time among factory workers.[37] It is apparent from this table that reducing the length of the working week produced a disproportionately greater reduction in amount of lost time. Thus with men factory workers the amount of lost time fell 43 per cent when

Table 10-3. The Relationship between Length of the Work Week and Amount of Lost Time among Factory Workers

	Length of the work week		Lost time	
	Hours	Relative time	Per cent	Relative time
Men	63¼	100	7.0	100
	62	98	5.7	81
	54	85	4.0	57
Women	62	100	6.4	100
	54	87	4.3	67
	44	71	3.1	48

the length of the work week was reduced by 15 per cent. Under many circumstances it would appear that the expected reduction in amount of production from shortening the hours of work is compensated for by the increased production effected by greater numbers of employees reporting for work.

REST PAUSES

Voluntary Rest Pauses. When systematic rest pauses are not set for the worker, it is ordinarily expected that he will take them voluntarily. These voluntary rest pauses appear to be a function of the nature of the work. being somewhat longer with heavy work than with light work. This fact is illustrated by some of the findings of Vernon, Bedford, and Warner which are given in Table 10-4.[40] These investigators also found that voluntary rest pauses are not taken by piece-rate workers to the extent that they are taken by workers paid on a time-rate basis.

The duration of rest pauses is also a function of the length of the workday. With a group of women factory workers, voluntary rest pauses were 28 per cent less in an 8-hr. than in a 10-hr. workday.[37] The voluntary rest pauses of silk winders, coil winders, and lathe operators were found to be 40 per cent less for those working a 41-hr. week than for those working a 49-hr. week.[38] In a tin-plate plant the voluntary rest pauses were 18 per cent less when the men worked a 6-hr. day than when they worked an 8-hr. day.[35] These findings suggest that a change from a very long working day to one of moderate length will produce a reduction in

Table 10-4. The Proportion of Time Consumed by Voluntary Rest Pauses in Several Occupations

Occupation	Operation	% of time per hr. spent in voluntary rest
Agriculture	Horse plowing	24
	Harrowing	14
	Rolling	14
Roadmaking	Picking and shoveling sand	21
	Picking up road surface	20
	Shoveling earth into wagons	19
	Shoveling sand into tubs	17
	Shoveling earth and wheeling barrows	13
Building	Bricklayer's laborers	19
Dock labor	Unloading timber	20
	Unloading sacks	18
Boot and shoe making	Consolidated lasting	6
	Pulling over	6
	Leveling	5
Rolling tin plates	Furnacemen	12
	First helpers	12
	Rollermen	12
	Behinders	12
	Second helpers	8
	Doublers	8

voluntary rest pauses disproportionately greater than the reduction in the working hours; whereas a change from a day of moderate length to a very short day, although reducing the length of voluntary rest pauses, will not produce a decrease proportionate to the reduction in working time. The proportionately longer time taken for voluntary rest pauses in very long and very short workdays, as compared with working days of moderate length, may partially account for the reduced output characteristic of the longer and shorter work periods.

Systematically Scheduled Rest Pauses. Vernon, Bedford, and Warner have pointed out that the rest pauses voluntarily taken by workers tend to have a certain regularity. Some of their observations on individual

Table 10-5. The Regularity of Voluntary Rest Pauses of Individual Workmen in Terms of Frequency and Length

Worker and job	Number of rest pauses in successive ½-hr. periods				Length of rest pause, min.	
	1	2	3	4	Average	Average deviation
A. Picking road surface	6	6	5	5	.8	.3
B. Picking road surface	5	4	5	5	.9	.3
C. Shoveling earth	5	5	7	4	.8	.3
D. Shoveling earth	6	5	5	4	.7	.2
E. Picking and shoveling	7	7	7	9	.6	.1
F. Picking and shoveling	8	9	7	9	.6	.2

workers are summarized in Table 10-5.[40] It will be noted in the table that regularity in the rest pauses is found in both the frequency and the length. A worker tends to take about the same number of rest pauses in equal successive time intervals, and the length of the pauses is relatively constant. These findings indicate that systematically scheduled rest pauses would be of considerable value, and, indeed, many studies have shown that this is true.

Systematically scheduled rest pauses have increased output in a variety of different occupations. In a boot and shoe factory, under an older system of work, two girls tended each machine and worked at it constantly.[20] Under a newer system a team of three girls was assigned to each machine and worked in shifts, so that two girls at a time tended the machine while the third girl rested. Each girl worked 40 min. per hour and rested 20 min. This change in schedule resulted in an increase in weekly output of 44 per cent, in spite of the fact that the total amount of time each girl worked was reduced by one-third. In addition the girls reported that they felt less tired under the new schedule. In an investigation of girls employed in labeling, the introduction of a 10-min. rest period in the middle of the morning was found to increase hourly production 20 per cent, whereas working time was reduced only 2 per cent.[39]

McGehee and Owen found that the introduction of a 12-min. rest pause in the middle of the morning and a similar pause in the middle of the afternoon increased the output of comptometer operators 29 per cent and decreased the length of voluntary rest pauses 60 per cent.[24] This study

is a valuable one because it shows that rest pauses in an office situation may prove as beneficial as rest pauses in factory or manual work situations.

It would appear that the increase in output reported in studies concerned with rest pauses should not be ascribed to the shortening of the working day, but rather should be attributed to the rest periods themselves. Evidence for this interpretation comes from the study of McGehee and Owen.[24] In this investigation the daily working time was lengthened by the amount of time devoted to rest pauses. Despite the longer day, the rest pauses increased the output and decreased the amount of time voluntarily taken for rest by the workers.

Factors Determining the Effectiveness of Rest Pauses. The effectiveness of rest pauses is a function of the time at which they are introduced and of their length.[47] The evidence indicates that they should be introduced when output is at a maximum, just before production begins to fall. Rest pauses introduced either earlier or later are not as effective in maintaining a high level of productivity. The optimal time between rest pauses varies from one job to another and will have to be determined separately for each. This is also true for the optimal length of the rest pause. The effectiveness of a rest pause will not be proportional to its length; rather there will be an optimal time, with longer or shorter times being less effective. If the pause is too brief it will simply interrupt the worker and not help in overcoming work decrement. On the other hand, if it is too long the worker will have to go through a fairly long warming-up period when work is resumed. In either case, output may even show a decline.

There is some evidence that less effective workers benefit more from rest pauses than do the better ones.[20,46] This is partially because the better workers have developed more effective methods of work that are less demanding on them. For example, Farmer found that among workers engaged in roughing spoons, the better workers used many fewer movements per item of work than did the poorer workers.[8]

When the work is purely automatic rest pauses are less effective than when the work involves considerable concentration.[47] This is to be expected since during automatic work the employee is free to carry on conversations, daydream, and engage in other activities that will reduce the monotonous character of the job.

It is ordinarily believed that the beneficial effects of rest pauses are purely physiological. However, as Pennock suggests, the resulting increase in output can be attributed in part to a change in the attitude of the employees toward the working situation.[28] He reports this finding from a study of a group of girls assembling telephone relays who were

subjected to a series of different conditions of work, among which were variations in the position and length of rest pauses. The production of these workers continued to increase regardless of the nature of the working conditions that were introduced. Even when rest pauses were completely eliminated, output continued to increase. During the experimental period, which lasted some two years, there was a marked improvement in the attitude of the girls toward the management. The relationship of confidence and friendliness that developed toward management seemed to account for the continuous rise in production. Another indication of the part played by attitude is found in Wyatt's report that workers actually anticipate rest pauses by showing an increased rate of work immediately preceding a rest pause.[47]

SHIFT WORK

One solution to the problem of obtaining maximal production with the minimum of overhead expense is to have employees work on successive shifts. There has been much discussion of the advantages and disadvantages of the shift system, and, although there are a number of scattered investigations of specific aspects of the problem, most conclusions have been reached without reference to facts or figures. It was not until 1934, when Vernon published a summary of the problems and statistical findings, that any adequate analysis was available.[38]

The confusion in the field can be illustrated by reference to a report of a committee of the Home Office of Great Britain made in 1920.[38] Witnesses to the committee were drawn from employees, labor union representatives, employers' organizations, physicians, social workers, and other similar groups. Although some of the witnesses said that shift work disturbed meals and hours of sleep, others denied it. The advantages of club life and evening classes were claimed to be interfered with by shift work, but it was pointed out that even under the most favorable conditions only a very small proportion of industrial workers took part in such activities. It was said that shift work interfered with the social life of the home, but in many cases this was found to be untrue.

Obviously, conclusions based on such opinions cannot be accepted. Knowledge concerning the effects of shift work must be based upon sound investigation and not opinion. Unfortunately, most studies either deal only with the two-shift system where employees working on the night shift put in fewer hours than day workers, or merely compare day work with night work. In actual practice several different types of shift systems are in operation.

Shift Work and Output. Vernon has summarized some of the findings relative to the effects of the two-shift system on output.[38] These results,

Table 10-6. A Comparison of the Output of Workers on a Two-shift System with That of Day Workers

Type of work	Hours of work per week		Relative output of two-shift workers, setting output of day workers at 100	
	Shift	Day	Hourly output	Weekly output
Coil winding	40¾	47	104	90
Coil winding	40¾	47	107	94
Coil winding	40¾	47	100	87
Engraving	40¾	47	107	92
Wire drawing	43½	47	98	90
Lathe operation	41¼	49 55	123	104 92
Lathe operation	41¼	49 55	118	100 89
Cable insulating	41½	49	135	113
Cable insulating	41½	49	125	104
Average	41¼	49¼	113	96

given in Table 10-6, indicate that the two-shift system, with its shorter working week, results in an hourly output that is considerably higher than that of the single shift, whereas weekly output is only slightly reduced. Clearly the combined output of both shifts on the two-shift system is vastly superior to that of the single day shift.

Studies in which day shift is compared with night shift in factory and manual work have yielded results contrary to what is ordinarily expected. Vernon has reviewed the studies of production on such jobs during the day and night shifts, and the findings are summarized in Table 10-7.[38] It will be observed that, when the workers are shifted from day to night work at regular intervals, output on the night shift is almost exactly the same as that on the day shift. The sex of the workers is not a relevant factor since the production of women workers on the night shift is maintained at a level as high as that of men. Furthermore, it appears that the nature of the work is not a determining factor. The work can be of a strenuous manual sort, as charging blast furnaces; highly repetitive work, as in the cartridge-case operations; completely handwork, as in sizing fuse bodies; or almost wholly machine work, as in boring top caps.

The data in Table 10-7 also indicate that when workers are kept continuously on the night shift their output is significantly reduced. This conclusion is supported by studies in which the output of a group of workers while on a continuous shift is compared with their output on a changing shift. Although the differences are not quite so striking, they are

Table 10-7. Comparison of Output on the Night Shift with That on the Day Shift

(Output of the day shift = 100)

Type of worker and work	Hourly output of night shift	Frequency of change from day to night shift
Cases involving periodic changes from day to night shifts		
Women, cartridge-case operations	101	Weekly; 52 hr. day shift, 55 hr. night shift
Women, cartridge-case operations	99	Same as above
Women, cartridge opr., 2d draw	102	Weekly; 51.7 hr. day shift, 55.9 hr. night shift
Women, cartridge opr., head trim.	98	Same as above
Women, cartridge opr., 2d cutoff	100	Same as above
Women, cartridge opr., reaming	102	Same as above
Women, cartridge operations	95	Weekly
Men, 6-in. shell operations	104	Every 4 weeks
Men, 9.2-in. shell operations	101	Weekly
Men, hand-charged blast furnace	99	Weekly
Women, turning fuse bodies	109	Every 2 weeks
Men, sizing fuse bodies	103	Every 2 weeks
Women, milling screw threads	103	Every 2 weeks
Youths, boring top caps	99	Every 2 weeks
Average	101	
Cases involving continuous night work compared with continuous day work		
Men, boring powder chambers	90	
Men, finishing, turning, 3-in. shells	98	
Men, rough turning 3-in. shells	94	
Women, cartridge opr., during winter	83	
Women, cartridge opr., during summer	88	
Women, bullet opr., during winter	90	
Women, bullet opr., during summer	100	
Average	92	

nevertheless in the same direction. With a group of women engaged in various cartridge operations, Vernon found that their hourly output for 4 weeks on a continuous day shift was only 1 per cent less than their output on a previous 6-week period and a subsequent 10-week period, when they were on a changing shift.[38] The hourly output of another group of cartridge workers during a 5-week period on a continuous night shift was found to be 4 per cent less than their output on previous and subsequent weeks when they were on a changing shift.

Shift Work and Lost Time. Vernon has compared both the voluntary and involuntary work stoppages of artificial-silk workers, coil winders, and lathe operators on a two-shift system with comparable groups of day workers.[38] The shift workers took voluntary rest pauses averaging 4.1 min. per hr. and involuntary pauses of 1.9 min.; the day workers took voluntary rest pauses of 6.8 min. per hr. and involuntary ones of 3.1 min. The higher hourly output of the shift workers therefore can be explained in part on the basis of less time spent in activities not reflected in production. The involuntary pauses were due to such activities as obtaining materials and adjusting equipment.

Vernon found that absences were more common among the shift workers than among the day workers. Although there were about 6 per cent more instances of tardiness among the day workers, tardiness among shift workers was usually much more serious. Avoidable absences also were less frequent during the night shift than during the day shift. With munitions workers, for example, such absences were 30 per cent less frequent with the night shift than with the day shift.

Shift Work and Health. The available statistics indicate that neither the two-shift system nor night-shift work has any deleterious effect upon health. In an artificial-silk mill, for example, it was found that the same proportion of workers on the two-shift system reported to the first-aid station for respiratory troubles, headaches, neuralgia, faintness, disorders of the digestive system, etc., as was found for day workers.[38] The effect of night work upon health cannot be ascertained in a direct manner due to a lack of information. However, with men employed at an engineering plant where every other week they changed over from the day to the night shift, it was found there was about 66 per cent less time lost for unavoidable causes on the night shift than on the day shift.[37]

PHYSICAL CONDITIONS OF THE WORK ENVIRONMENT

With the stress that is today put upon the effects of the social environment upon the worker, the importance of the physical environment is likely to be overlooked. There is no question that the group in which the worker finds himself has a profound influence upon his behavior. Never-

theless it is equally true that he is not insensitive to the physical conditions of his work environment. Problems connected with the social environment will be dealt with in a later chapter. In this section the more important aspects of the physical environment will be discussed.

Possibly the most important physical aspect of the working environment concerns the illumination of the working space. A survey of 21 industrial and office activities showed that the eyes are engaged in important work during 70 per cent of the working day.[42] It is therefore necessary to consider the effects of illumination upon the worker and his performance. A second aspect of the working environment needing attention is ventilation. There are a number of incorrect notions concerning just what constitutes adequate ventilation. It is known that feelings of discomfort have little or nothing to do with the "staleness" of the air, but rather are a function of the capacity of the atmosphere to cool the body. A third environmental factor is noise. Noise can operate as a distractor, and as such causes a reduction in output. The worker, however, can adapt to many noises, and when he does his effectiveness usually will be maintained at a normal level.

ILLUMINATION

Considerable evidence has been gathered that supports the validity of the following two important principles of effective illumination: [42]

1. The intensity of the illumination should be sufficient for the task.
2. The visual field should be as uniformly illuminated as possible with any bright spots placed as close as possible to the fixation point of the eyes.

Intensity of Illumination. Evidence that many industrial plants are inadequately illuminated comes from the fact that increasing the intensity of illumination produces increased output. Typical results obtained by Luckiesh and Moss are shown in Table 10-8.[22] Similarly, increases in intensity of illumination result in an increase in accuracy particularly with fine work such as typesetting.[43] Brighter light therefore results in faster and more effective work.[21] Other investigations reveal a similar improvement with increased intensity of illumination.[44]

Although increases in intensity of illumination may be expected to produce increases in output, a point is reached beyond which further increases in illumination will not be profitable or bring about further benefits to the employee. These points of optimal benefits have been termed critical levels of illumination. As might be expected they vary with the particular work situation and the kind of work. For reading of legible type the critical level of illumination is about 3 to 4 foot-candles, for arithmetic computations less than 10 foot-candles, for sorting mail 8 to 10

Table 10-8. The Effects upon Output of Increasing the Intensity of Illumination

Type of work	Old intensity, foot-candles	New intensity, foot-candles	% increase in production
Pulley finishing	0.4	4.8	35
Metal bearings	4.6	12.7	15
Steel machining	3.0	11.5	10
Carburetor assembling	2.1	12.3	12
Iron manufacturing	0.7	13.5	12
Buffing shell sockets	3.8	11.4	9
Letter sorting	3.6	8.0	4

foot-candles.[33] For more exacting types of work the critical levels are considerably higher, being, for example, 20 to 22 foot-candles for setting type and 30 foot-candles for threading a needle. Since values such as the above are only averages, a margin of safety ordinarily is provided to allow for individual variation. These adjusted levels become the standards of illumination. Although it might be expected that standards of illumination could be easily and accurately set, there actually is considerable disagreement concerning them.[4] This situation arises in part due to a lack of adequate experimental data concerning critical levels of illumination and in part from disagreement concerning necessary margins of safety.

Ferree and Rand have shown that the intensity of illumination required for effective work not only varies with the work situation and type of work but also with the worker.[13] For example, they found that with younger workers a high level of visual acuity is achieved at 10 foot-candles of illumination, and further increases in intensity produce only negligible increases in acuity. With middle-aged workers, however, such an intensity is inadequate, and it is only with further increases that maximum acuity is reached. In the case of older workers a low level of acuity is characteristic of nearly all intensities, and intensities beyond 20 foot-candles do not appreciably increase visual acuity. Under conditions of poor illumination differences among workers in productivity are magnified.[43] The output of persons with poor vision suffers proportionally more from poor illumination than that of workers with good vision. Individual differences in reaction to illumination are also exemplified by the wide variations in the lowest intensity of natural illumination which workers will accept before seeking to supplement the illumination by turning on artificial light.[43]

Improved output cannot be expected to follow immediately upon changing conditions of illumination. The employees must be allowed time to adjust themselves to the new working conditions.[34] In one case, with key-punch operators, production showed no marked increase until about 2 months after the intensity of illumination was changed from 8 to 60 foot-candles, and errors showed no marked reduction until after about 4 months.[34] The final level of output, an increase of about 45 per cent, and of accuracy, a reduction of about 70 per cent in errors, was not reached until nearly 2 years after the change was made. Although these times may be longer than usually is expected for a change in illumination to become manifest in output, the findings of this study emphasize the necessity for permitting the employees to work under a new situation for an adequate time interval before final evaluation of any change is made.

All the increase in output resulting from increased intensity of illumination cannot be attributed to increased visual acuity. In one investigation it was found that output was 8 per cent greater with work carried out purely by touch when the subjects wore goggles equipped with frosted glass that permitted light to reach the eyes but prevented detailed vision, than when they worked in the dark.[22] Apparently light has a facilitating effect upon workers. These findings may be due to the decrease in feelings of tiredness and the increase in feelings of well-being of the workers which have been reported by them when conditions of illumination were improved.

Uniformity of Illumination. Uniformity of illumination also shows its effects upon production. The strong tendency of the eyes to fixate the brightest part of the visual field is evidence that the visual field should be uniformly illuminated.[23] With inspectors, for example, it was found that when illumination was made more uniform and bright spots were reduced output increased by 9 per cent.[49] In actual practice it is next to impossible to have the total visual field absolutely uniform because of the variation in reflective powers of different materials and colors. The solution is to have any bright spots as close to the work as possible. In addition it is sometimes possible to have the illumination at the immediate workplace greater than that in the surrounding work environment.

Ferree and Rand have studied conditions of illumination during reading in relation to visual efficiency.[9] These investigators consider the ratio of blurred to clear time as a measure of what they term visual efficiency. Daylight gives the most uniform illumination, indirect artificial illumination the next most uniform, semi-indirect the next, and direct the least uniform. The percentage loss in visual efficiency during reading found by Ferree and Rand for these four types of illuminating conditions is given in Table 10-9. It will be noted that, with continuous reading, day-

Table 10-9. The Loss in Visual Efficiency during Reading under Four Types of Illumination Differing in Degree of Uniformity

Type of illumination	% lost after 2 hr.	% lost after 3 hr.
Daylight	5	6
Indirect	10	9
Semi-indirect	34	72
Direct	37	81

light produces the least loss in visual efficiency, indirect lighting produces more, semi-indirect still more, and direct produces the greatest loss.

The effect of the number of lighting units in the field of vision has also been investigated by Ferree and Rand.[10] The more fixtures there are in the field of vision, the more apt the eyes are to wander from the work in hand. The percentage loss in visual efficiency after 3 hr. reading, found by Ferree and Rand for semi-indirect and direct lighting with varying numbers of lighting units in the field of vision, is shown in Table 10-10.

Table 10-10. The Loss in Visual Efficiency during Reading When Various Numbers of Lighting Fixtures Are Placed in the Field of Vision

Number of fixtures	Type of lighting	
	Semi-indirect, loss, %	Direct, loss, %
0	30	40
2	40	47
4	50	67
6	78	84

It will be observed that, as the number of fixtures in the visual field increases, the percentage loss in visual efficiency increases. As would be expected, the loss from semi-indirect lighting is not so great as that from direct. Clearly the intensity of illumination is not the sole criterion of effective illumination; the character of the illumination is also important.

Color of Illumination. Even when the intensity of artificial light is equal to that of daylight illumination, it is inferior. The reason is that ordinary artificial light differs from daylight in uniformity and in color.[12] The importance of uniformity has already been discussed and the importance of color may be illustrated by a study made by Pierce and Weinland.[29] These investigators set up an experimental room, the walls of which were painted a flat white that could be illuminated by different colored lights, all of the same intensity. Subjects worked at a repetitive manual task in this experimental room under different colored illuminations. The results, shown in Table 10-11, indicate that none of the colors

Table 10-11. The Production in a Repetitive Manual Task as Related to the Color of the Illumination

Color of illumination	*Relative production*
White	100
Yellow	93
Green	92
Blue	78
Red	76
Orange-amber	76
Yellow-amber	54

tested is as effective as white light. The finding that yellow is the most effective of the colors has been corroborated by other studies.[11] Only when the uniformity and color of artificial light approximate those of sunlight are equal intensities of the two equally effective.

VENTILATION

A consideration of the problem of ventilation brings to mind matters connected with noxious and toxic conditions of the air, together with the effects of rarified atmospheres. While it is true that these are important problems, they are highly specific and have no bearing on most industrial jobs. The discussion of ventilation in this section will deal with the types of ventilation problems that are likely to arise in most normal working situations.

Feelings of Comfort and Oxygen Supply. It is a common belief that the bad effects resulting from inadequate ventilation are due to the reduced oxygen content of the air and to the excessive amounts of carbon dioxide present in it. This notion was completely disproved many years ago by Paul.[16] Paul's subjects sat in an airtight chamber breathing the same air over and over again. After about 5 hr. they reported the usual symptoms of discomfort. When the symptoms were at their height the

subject stuck his head into a separate chamber where he could breathe fresh air, but his body remained in the "stale" air. If the effects of bad ventilation were dependent upon the amount of oxygen and carbon dioxide in the air being inspired, the subject should have reported that his feelings of discomfort disappeared after the change. His discomfort, however, was not diminished. Paul then did a control experiment in which the subject's body remained in the outside room under normal air conditions while his head was in the experimental chamber, thus forcing him to breathe the "stale" air. In this instance the subject reported no symptoms of discomfort. Paul's findings received adequate corroboration in later studies made by the New York State Commission on Ventilation.[27] Clearly feelings of distress resulting from poor ventilation in the ordinary working situation cannot be attributed to the lack of oxygen and the accumulation of carbon dioxide in the air resulting from the respiration of its occupants.

The unimportance of the oxygen and carbon dioxide content of the air also is indicated by the fact that in most poorly ventilated schools and factories the oxygen content of the air seldom falls below 19 per cent, and the carbon dioxide content seldom rises above 0.3 per cent. In order to produce adverse effects the oxygen content must be reduced below 14 per cent, and the carbon dioxide content must reach 0.4 per cent.[30]

Ventilation and the Heat Regulation of the Body. The function of ventilation is to aid in the heat regulation of the body. The principal physiological system by means of which heat is reduced and comfort maintained is the evaporating mechanism of the skin, which produces cooling effects through perspiration. The rapidity of this evaporation is a function of the cooling power of the air, which, in turn, is dependent upon the temperature of the air, the amount of water vapor it contains (humidity), and its rate of movement. High temperature reduces the radiation and convection of heat from the body, high humidity reduces the rate of evaporation, and stagnant air impairs both convection and evaporation. By reducing either temperature or humidity, or by increasing the movement or circulation of the air, ventilating conditions are improved.

No individual can be comfortable and work effectively if the heat produced by the activity of his body is not dissipated. During physical activity the body produces considerable heat, and the indication is that even during "mental" activity some heat is produced over the amount given off during the resting state. If a person is kept in an airtight chamber, breathing the same air over and over, the feelings of discomfort can be eliminated and an effective level of production maintained by turning on a fan. The increased speed of movement of the air increases its cooling power.[27]

The Comfort Zone in Working Environments. In the working situation neither high nor low temperatures are comfortable. Bedford obtained workers' reports of comfort for various temperatures in a number of different factories.[2] The percentage of workers reporting the temperature as comfortable for various degrees of warmth and cold are given in Table 10-12. Apparently the optimal temperature for comfort is about

Table 10-12. The Percentage of People Reporting Various Temperatures as Comfortable

Temperature, °F.	*% of people reporting comfort*
54 to 55	28
56 to 57	46
58 to 59	58
60 to 61	77
62 to 63	84
64 to 65	78
66 to 67	69
68 to 69	60
70 to 71	45
72 to 75	45

62 or 63 degrees. It is important to note, however, that this optimal temperature was not reported as comfortable by all of the workers. Similarly, very high levels of humidity may be reported as uncomfortable, depending upon the movement of the air. The three variables of temperature, humidity, and rate of air movement should not be considered separately. With a constant temperature, varying the humidity will change feelings of comfort, and with both temperature and humidity constant, feelings of comfort will vary with the speed of movement of the air.[50]

The Effects of Ventilation on Production. Numerous studies have shown that production is a function of the goodness of the ventilation. A typical example of the increase in output resulting from improvement of ventilation under difficult conditions is reported by Wyatt, Fraser, and Stock.[48] These investigators studied output in weaving mills in which humidity must necessarily be high in order to prevent thread breakage. The tensile strength of thread is greatest in a moist atmosphere. To maintain high humidity steam was shot into the working rooms and this, of course, maintained the temperature at a very high level. As judged by the workers these atmospheric conditions never reached the minimum standard of comfort. The problem of discomfort was solved by introducing fans. When the fans were in operation output increased and breakage decreased, the improvement being greatest on those days when the

temperature and humidity were unusually high. In addition the reports of the workers indicated that bodily comfort was increased.

NOISE

Many controversies have been waged concerning the detrimental effects of noise on worker performance. As with most factors known to influence behavior significantly, no simple generalization can be made to encompass all of the facts. Noise greatly affects the performance of some workers and has little influence on that of others. Noise affects some types of work much more than others. Noise may show a detrimental effect when the worker is first adjusting to the job, but in time the effect may completely disappear.

Noise as a Distractor. Under certain conditions noise serves as a distractor to reduce the effectiveness of the worker. Noises that are irritating to the listener because of their intensity or their shrillness are frequently found to cause decreases in output. Interrupted noises are more disturbing than continuous noises.[5] During the learning stage of a new task workers often describe extraneous noises as distracting and unpleasant.[31]

The mere presence of a noise does not necessarily distract the worker. In a study of typists turning out form letters, Kornhauser found that only 3.2 per cent more lines were written but 1.7 per cent more errors were made under quiet conditions than under noisy conditions.[17] Actually, under the quiet conditions 23 per cent more letters had to be discarded because of faulty work. The typists reported that they felt they worked no harder in the noisy than in the quiet situation. Weston and Adams found that, in the noisy weaving mill, fitting workers with ear defenders only increased output 1 per cent.[45] Although the ear defenders did not completely eliminate the noise, they reduced it to a considerable extent.

Accommodating to Noise as a Distractor. Evidence indicates that if the worker remains in the noisy situation which has distracted him, he may eventually learn to adapt to it so that he regains any loss suffered in output.[31] This fact is often interpreted to mean that continued exposure to noise does not have any adverse effect upon the worker. Such a generalized statement is not justified.

Subjective discomfort from a noise often diminishes or completely disappears as the worker continues on the job. He comes to accept the noise, considering it as part of the natural background of the task. Cassel and Dallenbach found evidence that when a noise is accepted as part of the background it has no distracting effect.[5] Thus the noise made by individuals in connection with their work is not likely to disturb them.[1]

Part of the adaptation of a worker to a distracting noise involves the

expenditure of extra energy in his responses. In an early study, Morgan presented subjects with a task in which they responded to a series of stimuli by striking keys.[26] Effort, as measured by the extent of pressure exerted upon the keys, was greater during the noisy conditions than during the quiet ones. Output during the noisy conditions, rather than being reduced, was actually increased a little. Measuring respiration by means of the Douglas-bag method, Laird compared the energy consumption of typists working under quiet and noisy conditions.[19] Under the latter conditions 39 per cent more energy was consumed than under the former. Freeman has reported that under continued noisy conditions the increased effort and energy expenditure which occurs at the beginning of exposure to noise gradually decreases and finally approaches the level characteristic of quiet conditions.[14]

Distraction by Noise as a Function of the Nature of the Task. As is to be expected a distracting noise does not affect all tasks equally. Pollock and Bartlett found that noise fails to show much distracting effect with tasks that are readily automatized.[31] This is particularly true with simple motor tasks. Similarly, with tasks that are both highly interesting and complicated, noise does not appear to disturb output. On the other hand, with tasks that do not demand complete attention, noise tends to be distracting. The noise is especially disturbing in these tasks when the individual comes to a difficult point in the performance. Pollock and Bartlett's results agree with those of Vernon and Warner, who report that the more the work involves mental activity the greater is the disturbance from noise.[41] The latter investigators report that with mental work the amount of subjective disturbance due to noise decreases from the beginning to the end of the day and from day to day.

Distraction by Noise as a Function of the Worker. The discussion thus far has been confined to the working conditions under which noises distract. Another problem in need of study concerns relationships between personality traits and susceptibility to distraction from noise. There is very little adequate information available on the subject. Culpin and Smith studied the relative frequency of symptoms of nervous temperament among workers who were distracted by noise and workers who were not distracted.[6] Workers who reported that they were distracted by noise displayed 76 per cent more of these symptoms than did workers who were not distracted. Workers who are tired, bored, maladjusted, or who find their work uninteresting or difficult frequently complain about the distracting effects of noise.[1] In many instances their complaints are simply rationalizations and they are merely using the noise as a means for excusing their lack of success to themselves and to others. In quiet working conditions their complaints would be directed against other features of their environment.

Cumulative Effects from Continuous Exposure to Noise. In considering the problem of noise it is not sufficient to study noise as a distractor; the cumulative effects of prolonged exposure to noise should also be examined. The problem can be considered in terms of two kinds of effects, the organic and the psychological. It has been clearly demonstrated that prolonged exposure to noise produces permanent deterioration in the hearing apparatus of the inner ear. Boilermaker's disease is a common example. As a further instance, airplane pilots have been found to show a loss in sensitivity to high tones as indicated by audiometer tests.[7] The loss is sometimes apparent early in the pilot's flying experience, in some instances appearing after the first few hundred hours in the air. In the beginning the defect is temporary, but after continued exposure it becomes permanent.

A permanent psychological effect might be expected to manifest itself in terms of certain personality disorders. It is commonly believed that such disorders do occur among workers who have been employed for a considerable time in such noisy industries as stamping and weaving mills. The evidence, however, is primarily based on casual observation, and no definite conclusions are warranted at the present time.

REFERENCES

1. Bartlett, F. C.: "The Problem of Noise," Cambridge University Press, 1934.
2. Bedford, T.: The warmth factor in comfort at work: A physiological study of heat and ventilation, *Ind. Health Research Bd.*, No. 76, 1936.
3. Bills, A. G., and C. W. Brown: The quantitative set, *J. Exp. Psychol.*, **12,** 301–323, 1929.
4. Bitterman, M. E.: Lighting and visual efficiency: the present status of research, *Illum. Eng.*, **43,** 906–922, 1948.
5. Cassel, E. E., and K. M. Dallenbach: The effects of auditory distraction upon sensory reaction, *Am. J. Psychol.*, **29,** 129–143, 1918.
6. Culpin, M., and M. Smith: The nervous temperament, *Ind. Health Research Bd.*, No. 61, 1930.
7. Dickson, E. D. D., A. W. G. Ewing, and T. S. Littler: The effects of aeroplane noise on the auditory acuity of aviators: some preliminary remarks, *J. Laryngol.*, **54,** 531–548, 1939.
8. Farmer, E.: Time and motion study, *Ind. Fatigue Research Bd.*, No. 14, 1923.
9. Ferree, C. E., and G. Rand: The power of the eye to sustain clear seeing under different conditions of lighting, *J. Educ. Psychol.*, **8,** 451–468, 1917.
10. Ferree, C. E., and G. Rand: Lighting in relation to the eye, *Proc. Am. Phil. Soc.*, **57,** 453, 1918.
11. Ferree, C. E., and G. Rand: Visibility of objects as affected by color and composition, *Personnel J.*, **9,** 108–124, 475–492, 1931.
12. Ferree, C. E., and G. Rand: The transition from day to night lighting, *Personnel J.*, **11,** 237–254, 1932.

13. Ferree, C. E., and G. Rand: Sight and intensity of light, *Personnel J.*, **16,** 18–25, 1935.
14. Freeman, G. L.: Changes in tension pattern and total energy expenditure during adaptation to "distracting" stimuli, *Am. J. Psychol.*, **52,** 354–360, 1939.
15. Goldmark, J., M. D. Hopkins, and P. S. Florence: Comparison of an eight-hour and a ten-hour plant, *U.S. Public Health Service, Public Health Bull.*, No. 106, 1920.
16. Jenkins, J. G.: "Psychology in Business and Industry," Wiley, 1935.
17. Kornhauser, A. W.: The effects of noise on office output, *Ind. Psychol.*, **2,** 621–622, 1927.
18. Kossoris, M. D., and R. F. Kohler: Hours of work and output, U.S. Dept. of Labor, Bull. No. 917, 1947.
19. Laird, D. A.: Experiments of the psychological cost of noise, *J. Nat. Inst. Ind. Psychol.*, **4,** 251–258, 1929.
20. Loveday, J., and S. H. Munroe: Preliminary notes on the boot and shoe industry, *Ind. Fatigue Research Bd.*, No. 10, 1920.
21. Luckiesh, M.: "Light, Vision and Seeing," Van Nostrand, 1944.
22. Luckiesh, M., and F. K. Moss: "Seeing," Williams & Wilkins, 1931.
23. Luckiesh, M., and F. K. Moss: Light, lighting and seeing, in Wampler, F. J. (ed.): "Principles and Practices of Industrial Medicine," Williams & Wilkins, 1943.
24. McGehee, W., and E. B. Owen: Authorized and unauthorized rest pauses in clerical work, *J. Appl. Psychol.*, **24,** 605–614, 1940.
25. Miles, G. H., and A. Angles: The influence of short time on speed of production, *J. Nat. Inst. Ind. Psychol.*, **2,** 300–302, 1925.
26. Morgan, J. J. B.: The overcoming of distraction and other resistances, *Arch. Psychol.*, **35,** 1916.
27. New York State Commission on Ventilation: "Ventilation," Dutton, 1923.
28. Pennock, G. A.: Industrial research at Hawthorne, *Personnel J.*, **8,** 296–313, 1930.
29. Pierce, D. H., and J. D. Weinland: The effect of color on workmen, *Personnel J.*, **13,** 34–38, 1934.
30. Poffenberger, E. B.: "Applied Psychology," Appleton-Century, 1929.
31. Pollock, K. G., and F. C. Bartlett: Psychological experiments on the effects of noise, *Ind. Health Research Bd.*, No. 65, 1932.
32. Rothe, H. F.: Output rates among butter wrappers: I. Work curves and their stability, *J. Appl. Psychol.*, **30,** 199–211, 1946.
33. Tinker, M. A.: Illumination standards for effective and comfortable vision, *J. Consult. Psychol.*, **3,** 11–20, 1939.
34. Tinker, M. A.: The effect of adaptation upon visual efficiency in illumination studies, *Am. J. Optom.*, **19,** 143–151, 1942.
35. Vernon, H. M.: "Final Report," Health of Munitions Workers' Committee, London, 1919.
36. Vernon, H. M.: The speed of adaptation to altered hours of work, *Ind. Fatigue Research Bd.*, No. 6, 1920.
37. Vernon, H. M.: "Industrial Fatigue and Efficiency," Dutton, 1921.
38. Vernon, H. M.: "The Shorter Working Week," Routledge, 1934.
39. Vernon, H. M., and T. Bedford: The influence of rest pauses on light industrial work, *Ind. Fatigue Research Bd.*, No. 25, 1924.

40. Vernon, H. M., T. Bedford, and C. G. Warner: Rest pauses in heavy and moderately heavy industrial work, *Ind. Fatigue Research Bd.*, No. 41, 1927.
41. Vernon, H. M., and C. G. Warner: Objective and subjective tests for noise, *Personnel J.*, **11,** 141–147, 1932.
42. Viteles, M. S.: "Industrial Psychology," Norton, 1932.
43. Weston, H. C.: "Sight, Light, and Efficiency," Lewis & Co., 1949.
44. Weston, H. C., and S. Adams: Tenth annual report, *Ind. Fatigue Research Bd.*, 1930.
45. Weston, H. C., and S. Adams: The effect of noise on the performance of weavers, *Ind. Health Research Bd.*, No. 65, 1932.
46. Wyatt, S.: Variations in efficiency of cotton weaving, *Ind. Fatigue Research Bd.*, No. 23, 1923.
47. Wyatt, S.: Rest pauses in industry, *Ind. Fatigue Research Bd.*, No. 42, 1927.
48. Wyatt, S., F. A. Fraser, and F. G. Stock: Fan ventilation in a humid weaving shed, *Ind. Fatigue Research Bd.*, No. 37, 1926.
49. Wyatt, S., and S. M. Langdon: Inspection processes in industry, *Ind. Health Research Bd.*, No. 63, 1932.
50. Yagloglou, C. P.: Modern ventilation principles and their application to sedentary and industrial life, *J. Personnel Research,* **3,** 390, 1924.

CHAPTER 11

Methods of Work and Design of Equipment

Most tasks can be performed in a variety of different ways. It is apparent that effectiveness of work will be a function of the method utilized. Thus a nail is driven home with greater dispatch if the hammer is held at the foot of the handle rather than at the head, since holding the hammer at the foot permits a more powerful blow. Similarly, a long-handled hammer will be more effective in driving spikes than one with a short handle. It is therefore important to consider problems arising out of methods of work and design of equipment.

The discussion to follow will be concerned principally with work of a manual nature such as is found in factory operations, and most of the examples will be drawn from jobs of this kind. In part this is due to the industrial setting in which most of the problems in work methods arose. In many other kinds of jobs, such as clerical jobs, manual activities also are very important, and procedures for systematically studying work methods and a review of studies of manual activities will be of significance in areas other than factory operations.

TIME-AND-MOTION ECONOMY

At the present time the development of methods of work in industrial organizations is largely in the hands of the industrial engineer. This is understandable since so many industrial processes require technical knowledge relative to the design and operation of equipment. The approach of the engineer to problems in work methods has been termed time-and-motion economy. This approach grew out of the early job analyses of Taylor [25] and Gilbreth.[16] The analyses of these investigators were primarily designed to reveal how the movements of the body could be improved to bring about an increased productiveness in work. Thus time-and-motion analyses are the basis for studies of time-and-motion economy. In order to develop better methods of work for a given job, a motion-cycle chart is developed as a result of a time-and-motion study. The chart is then analyzed, unnecessary movements are noted and elim-

inated, and new movements or new combinations of old effective movements are suggested.

The Method of Time-and-motion Economy. It is beyond the scope of the present discussion to consider in any great detail the technique of time-and-motion economy. The large number of texts in the field testifies to its extent and manifold mechanics.[2] The method will be considered here only in its barest outline since we are only interested in its contribution to the problem of achieving the most effective motions.

As an example of the manner in which the economy is applied, a hypothetical case of a simple nut-and-bolt assembly will be described. In Table 11-1 are given fictitious data for the movements and times of

Table 11-1. A Hypothetical Example of the Application of Time-and-motion Economy to a Simple Nut-and-bolt Assembly Task

Worker A			Worker B			Best method		
Left hand	Time in sec.	Right hand	Left hand	Time in sec.	Right hand	Left hand	Time in sec.	Right hand
Reach for and select nut	1.1	Reach for and select bolt	Reach for and select bolt	0.7	Rest	Reach for and select bolt	1.1	Reach for and select nut
Carry nut to working area	0.2	Carry bolt to working area	Carry bolt to working area	0.7	Reach for and select nut	Carry bolt to working area	0.2	Carry nut to working area
Hold nut	3.2	Screw bolt into nut	Hold bolt	0.2	Carry nut to working area	Hold bolt	2.2	Screw nut onto bolt
Rest	0.5	Carry assembly to dump pile	Hold bolt	2.2	Screw nut onto bolt	Drop assembly to box under hands	0.1	Rest
			Carry assembly to dump pile	0.6	Rest			
Repeat cycle			Repeat cycle			Repeat cycle		
Total time 5.0			Total time 4.4			Total time 3.6		

movements for two workers. Although worker B takes less time to complete a cycle of work, it is apparent from an examination of the movements of the two workers that worker A is in some respects superior to B. Worker A takes only 1.3 sec. to pick up the two parts and bring them to the working position for assembly, whereas worker B requires 1.6 sec. The difference lies in the fact that worker A makes simultaneous movements with his hands, while the movements of B are consecutive. On the other hand, worker B is more effective than A in putting the nut and bolt

together, requiring only 2.2 sec. as compared with 3.2 sec. for A. Here it would appear to be quicker to turn the nut onto the bolt than to turn the bolt into the nut. For a best method of work, A's procedure for picking up the parts is more effective than B's, and for assembling, B's method is more effective than A's. Finally, both workers place their completed assemblies away from the central working position. This is unnecessary, and in a best method of work the assemblies simply could be dropped into a box underneath the hands.

Putting together the superior movements of A and of B and eliminating the unnecessary movements yields a method of making the assembly which requires much less time than the method employed by either worker. Production would be 22 per cent greater than that achieved by B's method and 39 per cent greater than that achieved by A's method. On the average, the new method presumably would increase production by about 30 per cent.

The Use of Time-and-motion Economy as a Basis for Increasing Effectiveness of Work. From this illustration it can be seen that time-and-motion economy involves the retention of those movements or combinations of movements in an older method of work that are performed quickly, and the elimination of extraneous movements and periods of inactivity. It is believed by exponents of the method that the resultant new method of work not only gives greater productivity, but that it also establishes the exact time needed to complete a unit of work. Thus production plans, rates of pay, etc., can be readjusted on the basis of the new work schedule. Essentially time-and-motion economy results in a description of hypothetical behavior that is used as basis for prediction in further planning.

EVALUATION OF TIME-AND-MOTION ECONOMY

Time-and-motion economy originally grew out of attempts to increase output and primarily has been in the hands of the industrial engineer. It is not surprising that the industrial psychologist, whose field of study is human behavior and who seeks to evaluate solutions to industrial problems in terms of their wider human implications, should find limitations in time-and-motion economy.[13] There is, of course, much about time-and-motion economy with which the psychologist is in accord. But these areas of agreement are principally concerned with administrative means for improving production, such as relieving skilled workers of those simple activities that can be performed just as well by unskilled persons, or eliminating unnecessary idle time. These aspects of time-and-motion economy are problems of production management rather than of psychology.

The psychologist's quarrel with time-and-motion economy is in terms of the basic postulates on which it rests. It follows that, if the basic postulates are false, then the method cannot be expected to give an exact prediction of the worker's behavior. It was noted in the chapter dealing with measurement of job proficiency that in time study certain allowances ordinarily are added to predicted times for completion of a unit of work, allowances being made for personal time, delays, etc. These allowances are established chiefly on an arbitrary basis. This fact alone is sufficient to demonstrate that the method is not exact.

The psychologist's concern is with time-and-motion economy as a technique for improving the effectiveness of work by revising the movements made by the worker in carrying out his job. From the point of view of industrial psychology, time-and-motion economy is open to criticism on four counts as follows: it is based upon an unsound atomistic conception of human behavior, it assumes a high degree of constancy of performance, it fails to take into account individual differences, and it uses far too narrow a criterion in gauging the effectiveness of work.

Atomistic Conception of Human Behavior. Time-and-motion economy is based upon the fallacious assumption that each individual movement of a total complex task can be considered as a separate and independent unit. Some time-and-motion analysts have fallen into the error of thinking of human behavior as being capable of analysis into distinct units like blocks which are summative in nature, and which may be arranged now into one pattern and now into another without affecting the character of any single element. For example, it is believed that if one movement is eliminated from a task, then the total time for completing the task is reduced in proportion to the time required to execute that movement as part of the entire sequence. This atomistic notion of human behavior is open to serious criticism.[26] The individual working at a task, however simple, knits all the part activities into a whole; therefore the changing of any part activity may be expected to change the pattern of the larger task.[7]

The inadequacy of this atomistic conception of movements may be demonstrated by the results of a very simple experiment. Subjects were presented with six keys arranged in the following spatial pattern:

A	D	F
B	C	E

They were instructed to tap the keys in a certain sequence, the time needed to make the movement between each pair of keys being recorded. Some of the movements were eliminated in a second sequence, and the times of movements again recorded. The recording of the times of the sequences was made after the subjects had mastered the task and

had reached their maximum speed of movement. Two examples of the results found with the key-tapping task are given in Table 11-2. In the

Table 11-2. Effects of the Elimination of Movements on Speed of Activity in Successive Tapping of Keys

Movement	Average time, sec.		
	First sequence	Second sequence	Ratio of second to first
Example I:			
AB	0.40	0.44	110
BC	0.28	0.30	107
CD	0.39	Eliminated	
DC	0.32	Eliminated	
CE	0.30	0.29	97
EF	0.39	0.41	105
FA	0.34	0.37	109
Total time	2.42	1.81	
Expected total time		1.71	
Example II:			
BC	0.23	0.21	91
CD	0.30	Eliminated	
DC	0.29	Eliminated	
CE	0.17	0.11	65
EA	0.31	0.30	97
Total time	1.30	0.62	
Expected total time		0.71	

first example the original sequence required 2.42 sec. The elimination of two movements that required 0.71 sec. was expected to reduce the total time to 1.71 sec. (2.42 minus 0.71 = 1.71). The actual time required to perform the second sequence was 1.81 sec. All but one of the movements retained in the second sequence required more time than they had required in the original sequence. Elimination of movements not only did not reduce the time as much as was expected but actually increased the times taken to perform most of the remaining individual movements. In the second example, opposite results were obtained. The elimination of some of the movements reduced the time much more than was

expected. The time required to perform one of the remaining movements in the new sequence was significantly reduced.

Barnes and Mundel have reported an experiment that bears on this point.[3] These investigators had subjects pick up pins and insert them one at a time into a hole, separate times being noted for the various parts of the operation. In all tests the hole was ¼ in. in diameter, but the aperture in four of the five experimental conditions varied in size and was funnel-shaped. The results, given in Table 11-3, indicate that the

Table 11-3. The Times Involved in Performing Various Parts of the Simple Task of Inserting a Pin into a Hole, in Relation to the Size of the Aperture of the Hole

Size of aperture, in.	Time to carry pin from supply to hole, sec.	Time to position pin for insertion, sec.	Time to insert and remove pin, sec.
1	.311	.281	.190
¾	.304	.327	.184
½	.347	.290	.179
⅜	.369	.295	.176
¼	.384	.485	.133

times required for the various parts of the task varied with the size of the aperture. Variation is to be expected in the time needed to position the pin for insertion as well as in the time required to insert it. However, it is not apparent why the time required simply to carry the pin from the point of supply to the hole should also vary with the size of the aperture. In any planning based upon time-and-motion economy this movement would be considered to remain constant, regardless of the movements that would precede or follow it.

It is clear that the time required for executing one part movement in a total sequence will vary with different characteristics in the working situation and with other part movements that may be only indirectly connected with it. It must be concluded that any task involving a series of movements will be integrated into a whole by the worker and that each component movement may interact upon and affect other component movements. The extent of this interaction may be great or small but it will exist to some degree and cannot be assumed to be absent.

Assumption of a High Degree of Consistency in Human Performance. The reader of discourses on time-and-motion economy frequently overlooks the assumption implicit in the method, that human performance in

any given activity displays a high degree of consistency. In an earlier chapter on measurement of job proficiency, it was pointed out in the discussion of reliability that the consistency of human performance varies markedly from one task to another. Behavior is never perfectly constant; even very simple activities are variable.

One point that is uncritically accepted in time-and-motion economy is that the statistical averages used are reliable. In a time-and-motion-economy study usually only a few cycles of the task are recorded, and the times for the different parts of the task under consideration are based upon averages obtained from these few cycles. The simple average as a statistical description of a set of data may or may not possess acceptable reliability. If the data are highly consistent, then their average will be a reliable description. On the other hand, if the values from which the average is computed vary markedly, then the average will be an unreliable description. Since it is merely assumed in the method that the performances being studied are highly consistent, it should not be further assumed that the averages are statistically reliable.

The degree of variability in the part movements of a simple task are illustrated in Table 11-4. Here are given the results of a time-and-motion

Table 11-4. The Times Required by Two Workers to Form a Link from a Piece of Wire, the Process Being Repeated 11 Times

Activity	Worker A				Worker B			
	Average time		Range of times		Average time		Range of times	
	Min.	%	Min.	%	Min.	%	Min.	%
Reaches for piece of wire	0.005	8	0.004-0.005	20	0.007	8	0.006-0.008	26
Selects and grasps piece of wire	0.004	7	0.002-0.009	175	0.005	6	0.002-0.009	140
Carries wire to bending apparatus	0.006	10	0.005-0.007	33	0.008	10	0.006-0.012	75
Inserts wire in apparatus	0.007	11	0.005-0.009	57	0.013	16	0.005-0.017	92
Forms link with apparatus	0.008	13	0.007-0.009	23	0.010	12	0.009-0.010	10
Operates apparatus lever	0.003	5	0.000	0	0.006	7	0.005-0.007	33
Takes out wire and inserts other end in apparatus	0.018	30	0.015-0.027	72	0.019	22	0.017-0.022	26
Forms link on other end of wire	0.007	11	0.000	0	0.010	12	0.009-0.010	10
Operates apparatus lever	0.003	5	0.000	0	0.006	7	0.005-0.006	17
Total	0.061	100	0.055-0.069	23	0.080	100	0.071-0.114	51

study made of two workers engaged in forming links out of pieces of wire.[1] Times were taken for each worker during 11 cycles of work. Of particular pertinence here is the variability of specific movements. In terms of the total time required to complete a cycle of work, the range

of times for worker A was 23 per cent of her average time and that for worker B, 51 per cent. For certain of the part activities the amount of variation in performance is exceedingly high. For most of the part activities of this task a statistical average is not sufficiently representative to be considered an accurate description of the activity. It should be apparent that some performances are so variable that an exact prediction of their times is impossible. Even when the averages for part activities are based on many cycles, there still will be many of these activities so variable that the average will not serve as an effective descriptive statistic.

Failure to Take Individual Differences into Account. The avowed purpose of time-and-motion economy is to discover the best method of work. The assumption is made that there is one sequence of movements that will be best for all workers. This notion completely denies the fact that individuals differ. Two workers may be equally effective in the performance of a task but nevertheless employ widely different movements in accomplishing it. There is no reason to suppose, as Farmer has pointed out, that a standard set of movements will be the best for all workers.[13] This is not to deny that optimal methods of work can be found, but it is one thing to think of a particular method of work as being best for *all* persons and quite a different thing to think of this method as being more effective than another for *most* persons.

Jenkins has suggested that the worker should not be adapted to the method of work but rather the method of work should be adapted to the worker.[22] Farmer has found that the study of the individual worker, with the aim of finding the best set of movements for him, is highly effective.[13] This suggests that it would be profitable to discover the kinds of methods that are optimal for different types of workers. Thus one method might be better for older workers and another for younger ones.

Narrowness of Basis Used for Evaluating Effectiveness of Work. Time-and-motion economy invariably evaluates the effectiveness of industrial work in terms of speed of production. When a number of movements are being compared, that movement which requires the least time is always considered to be the best. However, as Farmer has pointed out, the quickest movement is not necessarily the best in all respects.[13] Although rapid movements result in higher levels of production, they may also increase feelings of tiredness to such an extent that after a time the worker may be unable to continue on the job and may even rebel. Farmer obtained the cooperation of an excellent worker who was highly motivated in her work and who volunteered to work as fast as she possibly could for a trial period. Her output increased 105 per cent for a few hours, but then she was so exhausted that she could no longer continue working.

While the industrial engineer will not knowingly substitute a quicker though less safe motion for a slower but more safe one, the very position

he holds in the industrial organization forces his attention to matters concerned with increases in production. By and large his efforts are gauged in terms of how much he is able to increase output, rather than in terms of other benefits that might accrue. Thus advantage is not ordinarily taken of other types of benefits that could accrue from a broader application of the methods of the industrial engineer.

The various ways in which effectiveness of work can be gauged can find appropriate application in evaluating methods of work. While there is no question but that the amount of production is an important way of evaluating methods of work, it is also true that worker satisfaction, accuracy of performance, feelings of tiredness, etc., are equally important indices. It will be well to heed Farmer's warning that as long as the investigation of better methods of work is left completely in the hands of those whose primary purpose is to increase production, the worker, justifiably or not, will tend to feel that he is being exploited. On the other hand, Farmer has demonstrated that when the worker realizes that the purpose of a time-and-motion study is the reduction of fatigue he will regard the matter in an entirely different light.

TIME-AND-MOTION STUDY AS A BASIS FOR INCREASING EFFECTIVENESS OF WORK

In view of the above-mentioned criticisms it is doubtful that time-and-motion economy can be justified as an exact method. If it is based upon tenuous assumptions, then any predictions of human behavior made on the basis of it necessarily will be in error. In the foregoing examples it was shown that such predictions are inaccurate. Not only are the means inadequate, but the ends also cannot be considered wholly justified. This is not to imply, however, that methods of work cannot be improved through systematic investigation. Starting with such a detailed description of the particular movements now made in performing a job as would be given by a time-and-motion study it is possible to approach systematically the problem of developing improved methods of work.

Difference in Aims of Time-and-motion Economy and Time-and-motion Study. As stated in an earlier chapter, time-and-motion study is a systematic procedure for obtaining a description of the motions involved in performing a particular task. Its purpose is solely to provide a picture of the manner in which a job is being performed. Time-and-motion economy uses the results obtained from time-and-motion study as a point of departure. The former begins where the latter leaves off. Time-and-motion economy uses the results of time-and-motion study as a basis for drawing conclusions concerning ways in which the speed of work can be increased.

In addition to improved methods of work time-and-motion economy also employs other means for increasing production. Farmer has pointed out that a careful examination of discussions of time-and-motion economy and illustrative applications of it will reveal that not infrequently changes other than those made in the methods of work are introduced.[13] Thus, on the basis of an analysis of the movements involved in a job, a new sequence of movements will be set up, and, it is important to note, new rates of pay are likely to be instituted and perhaps a change in duties required of the worker. Any increases in productivity cannot then be ascribed simply to the new methods of work since other conditions also have been changed. Increases in production resulting from the application of time-and-motion economy therefore cannot be used to justify the new methods of work. If increases in production do occur, justification for the entire procedure used is obtained but not justification for any one aspect of it.

Time-and-motion Study as a Basis for Formulating Hypotheses Concerning Better Methods. Recognizing the fact that time-and-motion study is simply the systematic description of the movements required by a job and, in and of itself, tells nothing of the ways in which a job can be improved places the whole problem on a sounder foundation. Time-and-motion analysis therefore does not need justification on the basis of how much it improves methods of work but stands independently on its own right as a descriptive method for investigating human work.

In some instances time-and-motion study may directly reveal ways in which work can be improved. The illustration given in Table 11-1 of a simple nut-and-bolt assembly is a case in point. Here the analysis showed that the workers were making a large movement in carrying the completed assembly to the outer borders of the working area when they more easily could have dropped the assembly into a container under the hands. The difficulty in drawing a conclusion such as this one from a study of the pattern of movements involved lies in the fact that there might be elements in the situation that are overlooked. Indeed this is approximating the method of time-and-motion economy. Therefore changes in methods of work arrived at from simple inspection of the motion-cycle chart must be made with caution and should not be accepted as final until their effectiveness has been demonstrated empirically.

For the most part, when the industrial engineer employs time-and-motion economy as a basis for the improvement of work methods, he really is setting up an hypothesis that the new method he suggests is better than the old method. This obviously is true since after the new method is set up, a test is usually run with one or more workers trained on the new method, and their production is compared with that hypothesized or forecast by the economy procedures.

In so far as improvement of methods of work is concerned, time-and-motion study will be most valuable as a means for indicating specific research projects that should be conducted in studying the effectiveness of types of movements or methods of work. This can be illustrated by the results of a time-and-motion study made of the operations of a clerk in posting entries on a ledger sheet by means of a bookkeeping machine. The motion-cycle chart constructed from motion pictures taken of a number of cycles of work indicated that the left hand was idle for a period amounting to some 20 per cent of the total time. During this period the right hand was engaged in inserting numbers into the keyboard, and there was nothing by way of handling papers or other activities that the left hand could be doing. It was suggested that some time might be saved by simultaneously inserting numbers into the keyboard with both hands. Since there was no experience with the speed and accuracy of bimanual insertion of numbers, a study of the feasibility and effectiveness of this type of operation was indicated. In this case the motion analysis directed the thinking of the investigators toward a specific research project related to a particular method of work.

Time-and-motion study is likely to raise more problems than it solves. Since the objectives with respect to improving methods of work are usually stated in general terms such as increasing production or reducing work decrement, time-and-motion analysis makes a significant contribution in discovering specific areas needing study and in specifying particular problems that are amenable to immediate attack.

PROBLEMS IN THE ANALYSIS AND EVALUATION OF MOVEMENTS

Before discussing effective methods of work it will be well to consider some factors that may affect the results of motion analyses. It is important to learn about the circumstances that may limit or reduce the accuracy of such studies and thereby severely restrict the conclusions that may be drawn from them.

Representativeness of Conditions under Which Movements Are Studied. It is apparent that an investigation of movements made under one set of conditions may give results at variance with those obtained under another set of conditions. Thus Gilbreth noted that bricklayers used different movements when laying bricks quickly than when laying them slowly and used still a different set of movements when they were demonstrating how to lay bricks.[17] Similarly it may well be expected that the set of movements used in a particular task will be different under conditions of inadequate illumination than under conditions of superior illumination. An investigation of motions therefore should be conducted

under exactly the same conditions as those which may be expected to characterize the actual job.

Perhaps the most variable condition will arise in connection with the factor of motivation. The performance of a worker who is being used as a subject in a motion study will be a direct function of his attitude toward the situation. If he fears that the results will be used as a basis for setting up higher rate requirements, he will work at a slower pace than he might under other circumstances.

Even if the worker is certain that the results of the study will not be detrimental to his interests, factors may still operate to produce non-representative behavior. The mere fact that he is being observed or photographed may influence his performance. For most workers, some time is required to adapt to the newness of the situation. Whether the period of adaptation is to be measured in terms of hours or of weeks is difficult to determine because of the number and complexity of the factors conditioning this adaptation. Certainly the time will not be a matter of minutes, and possibly not even hours. It should be whatever time is required by the worker to attain a performance that is truly representative.

Representativeness of Individuals Used in Motion-study Investigations. The workers used in time-and-motion studies are likely to be unrepresentative of the total group of workers on the job. In many cases a foreman or a superior worker is used as the subject on whom the analysis is made. Such persons may differ from the general run of workers not only in terms of their speed of work but also in terms of the number and kinds of movements that they utilize.[13] If a scientific solution is sought, it is necessary to make sure that the motion analysis is made from the performance of workers who are truly representative of all levels of proficiency. Otherwise the resulting analysis will be severely limited in its application.

A similar problem exists in connection with laboratory studies of the effectiveness of different types of movements. In certain instances college students, or the investigators themselves, may serve as subjects. Such individuals will differ from employed workers not only with respect to motivation but also with respect to their pattern of abilities and previous work experiences. Differences such as these may well invalidate findings about the effectiveness of different types of movements, especially when these findings are generalized to the actual industrial situation.

Representativeness of Conditions in Laboratory Investigations. As has been seen, a time-and-motion study is likely to suggest specific research projects about particular methods of work. This fact, together with a growing inclination to substitute the findings of experimental investigation for the intuitive conclusions developed by time-and-motion economy,

has led to increasing numbers of laboratory investigations concerned with the effectiveness of different types of movements. These laboratory investigations received an impetus during the Second World War when a variety of problems arose in connection with new types of equipment and new types of jobs.

Certain of the laboratory investigations, however, have not escaped the error characteristic of time-and-motion economy. For example, some investigators accepted the dictum that a movement performed in isolation retains the same characteristics as when it is performed as one part of a total sequence of acts. Thus in a laboratory experiment a comparison of the effectiveness of curved movements of the arm with movements along straight lines might be made simply by having the subjects move their arms in the specified directions and at supposed maximum speeds of movement. If the speed of the curved movement is found to be higher than that for the straight-line movement, the former would be considered to be more effective under all circumstances, whether the movement was involved in picking up items for an assembly operation, in replacing tools on a bench, or in moving controls while operating a machine. It is apparent from the criticisms of time-and-motion economy already presented that such an interpretation cannot be considered scientifically justified.

The fact that the evaluation of a movement performed in isolation might be different from an evaluation of the same movement executed as part of a total act can be illustrated by a simple study. In this study a comparison was made between the effectiveness of simultaneous movements of both arms in opposite and symmetrical directions and simultaneous movements of the arms in the same direction. The subject was seated at a workbench. For symmetrical movements one box was arranged 1½ ft. to the right of the center of the body and another box the same distance to the left. When asymmetrical movements were made, both boxes were arranged on the right. In the first part of the experiment the subject simply moved his hands from the middle of the workbench to the boxes and back again. In the second part he reached to the boxes, selected a nut from one and a bolt from the other, carried them back to the center of the workbench, assembled them, and dropped the completed assembly to the workbench. Thus in the first part of the experiment the simple movements of the hands in symmetrical and asymmetrical directions were studied, and in the second part these movements were included as part of a sequence of movements.

In the simple movement of the arms the asymmetrical movements were made at the rate of 60.3 per minute and the symmetrical at 74.4 per minute. From these results it would generally be concluded that symmetrical movements are more effective than asymmetrical. In the second

part of the experiment, however, it was found that the number of assemblies made per minute when asymmetrical movements were employed was 6.6 per minute, and for symmetrical movements, 6.3. In this case the difference in the symmetry of the movements did not appear to produce significant differences in effectiveness.

It is apparent from the foregoing discussion that to obtain valid results from a laboratory investigation, movements should not be studied in isolation. Rather they should be included as part of an entire task in order that a correct evaluation can be made. In the problem on the relative effectiveness of symmetrical and asymmetrical movements, it would be necessary to compare the two types of movements when the arms were engaged in picking up and depositing various kinds of objects in executing a wide variety of different tasks before a statement would be justified concerning the general effectiveness of the two kinds of movement.

TYPES OF MOVEMENTS

Before considering the experimental findings with regard to the effectiveness of different methods of work, it will be well to have a clear understanding of different types of movements. The various movements that a worker makes can be classified in a variety of different ways. Three different ways in which movements are classified are: their purpose, their skill requirements, and their physiological nature.

Purpose of the Movements. As was pointed out in the earlier discussion of job analysis, time-and-motion study employs a classification of movements into types called *therbligs*.[2] Each therblig constitutes a basic element of movement and may be used to symbolize a number of different movements which have the same purpose. Thus, whether the worker is carrying in his hand a tool, a bolt, or a fixture of some sort, the purpose of the movement is the same, to convey an object from one place to another. All such movements are grouped together in the category of "transport loaded" and are given the same therblig notation. The therblig classification does not describe the nature of the movement, such as whether it is fast or slow, easy or difficult. This type of classification is important principally in such problems as changing the sequence of movements and eliminating unnecessary movements.

Skill Requirements of the Movements. One of the major problems in many jobs is to simplify the tasks in terms of the skill requirements. Job simplification has generally taken the form of reducing the number or variety of movements a worker needs to make in performing the job. The assumption is made that by this means the worker may utilize a lower degree of skill and will require less training. It does not necessarily follow, however, that the reduction of the number and variety of motions

will invariably make the task simpler for the worker. Indeed in some cases it might have just the opposite effect. It has been found that the more skillful workers generally employ fewer movements than the less skillful. This suggests the possibility that, in terms of skill requirements for a given task, reduction in the number of movements might actually require greater skill on the part of the worker.

A more fruitful approach to the problem of job simplification appears to be the classification of movements into types according to the nature of their skill requirements. Fairchild attempted to do this by classifying movements in terms of basic skills.[12] Each of the types of skills was defined; for example, *motion saving* refers to the worker's ability to direct a movement with control and without error, and *adaptation to machine* refers to the worker's knowledge and habits that enable him to regulate the machine and operate it and use it in accordance with its capacities.

In making the skill analysis a time-and-motion study is conducted, and each appropriate skill factor is noted in the motion-cycle chart along with the therblig to which it refers. This gives an analysis of the job not only in terms of the movements made but also in terms of the skill requirements. For any given job, dividing the number of skill factors by the number of therbligs gives a ratio that reflects the skill requirements. Ratios calculated for different workers or for different jobs permit rough comparisons in terms of skill requirements.

Fairchild's technique of skill analysis is cited here merely by way of illustration of a systematic approach. Many investigators will not agree either with her concept of skill or with her method of analysis. Certainly doubt can be cast upon the use of the skill-therblig average, which involves items that differ qualitatively and are not amenable to such arithmetic treatment. Nevertheless her method opens an attack on the problem from a different point of view, and further work along these general lines should be fruitful.

Physiological Nature of the Movements. Two types of voluntary movements of parts of the body may be distinguished: fixed or controlled movements and ballistic movements. In fixed or controlled movements, opposed or antagonistic muscles operate against each other. When the contraction of one set of muscles is stronger than that of the other, movement of the body part results. The thumb-and-finger method of writing is an example of this type of movement. In ballistic movements the movement of the part is initiated by the contraction of the appropriate muscle group, and the motion continues owing to the momentum that is generated. The contraction of the active muscle group after the initiation of the movement, and the contraction of its antagonistic muscles throughout the whole movement, are at a minimum. The arm-movement method of writing is an example of a ballistic movement. A ballistic movement is

stopped by the contraction of the antagonistic muscles, by the dissipation of the momentum of the movement, or by the part of the body involved striking an object. In order to employ momentum effectively in the movement of parts of the body, ballistic rather than fixed or controlled movements should be used.

Hartson has given considerable attention to the relative effectiveness of fixed and ballistic movements in skilled activities and finds the following advantages for the ballistic type of movement: [18]

1. Less decrement. Since ballistic movements are usually longer than fixed, it is ordinarily thought that there will be more work decrement. However, the reverse appears to be true, which might be expected inasmuch as ballistic movements presumably involve less muscular activity.
2. Less deterioration in the quality of movement with continuous activity. If a series of acts employing fixed movements is continued for a long time, the movements are soon characterized by decreased accuracy. In finger writing, for example, there is a rapid deterioration in the quality of the writing.
3. More powerful. Since ballistic movements use momentum and the force of the body weight, they are more powerful than fixed movements.
4. More rapid. The greater rapidity of ballistic movements has been demonstrated in a number of activities such as tapping, typing, and packing operations.
5. Less likelihood of cramp. Forms of "occupational neuritis" such as writer's cramp, telegrapher's wrist, and baseball pitcher's glass arm are much less likely to occur when ballistic movements are employed.
6. More accurate when speed is a consideration. With fine movements requiring considerable accuracy, fixed movements will be more accurate. On the other hand, with larger movements in which speed is important, ballistic movements are found to be more accurate.

In certain situations utilization of momentum is not desirable, since the worker must use his muscles to counteract it. If the momentum could be avoided, the worker would not have to expend energy to overcome its effects. The momentum of an object is equal to its weight (mass) multiplied by its velocity. Therefore, if either weight or speed of movement were reduced, the momentum would be reduced. The objects moved by the worker are materials, tools and equipment, and parts of his body. The weight of materials usually cannot be reduced, but it is often possible to reduce the weight of tools and equipment. In order to reduce the weight of the parts of the body that are moved it is sometimes suggested that finger movements be used rather than wrist movements, and wrist movements rather than arm movements. However, there is some evidence against this preferential hierarchy of movements.[24] In certain instances finger movements are likely to be less accurate and show more decrement than arm movements. Another way of reducing momentum is to have the arms and hands move through the shortest distance possible; hence, rearrangement of the workplace so that materials are found at points close to where they are to be manipulated will minimize the part played by momentum.

EFFECTIVENESS OF HAND AND ARM MOVEMENTS

Treatises on time-and-motion economy present a variety of principles regarding the effectiveness of various movements of the hands and arms. To a large extent these principles are based upon studies in which more than one variable was changed in the situation in which a given method of work was being investigated. Under such circumstances it is difficult to ascertain the extent to which change in any particular variable contributed to the resulting increase in production. In the following sections, therefore, consideration will be given only to evaluations of different methods of work that are based upon experimental investigations. The intent will not be to review all the many investigations on the effectiveness of movements, but rather to consider certain ones that appear to be of most importance in so far as industrial work is concerned.

One-hand Movements versus Simultaneous Bimanual Movements. It is a basic principle of time-and-motion economy that both hands should be working at the same time. In part this is done to save operation time, but in part it is based upon the belief that simultaneous motions of both arms are more natural and easier than individual movements of either arm. As evidence of the greater effectiveness of the bimanual method of work, investigations such as the following one by Barnes are offered.[2] In an older method for a bolt-and-washer assembly the worker picked up a bolt with her left hand, and then with her right hand she picked up and placed on the bolt in succession a lock washer, a plain steel washer, and a rubber washer. The bins from which the worker picked up the parts and into which she dropped the completed assembly were arranged one next to the other in a straight line on the worktable in front of her.

In a newer method of work the bins containing the parts were arranged in a semicircle on the table in front of the worker. To the central bin, which contained bolts, the worker reached with both hands and picked up two bolts, which she placed pointing upward in two countersunk holes in the table before her. Then she reached with both hands to the bins on either side of the central bin, both of these bins containing lock washers. Selecting a lock washer with each hand, she carried one to one bolt and the other to the second and assembled them. The next two bins from the center contained the steel washers, and the outermost two bins, the rubber washers. In turn the worker took two of each kind and assembled them on the bolts. The two hands moved simultaneously, neither being idle while the other was in operation, beginning and ending their movements at the same time and in opposite and symmetrical (mirror image) directions.

That the newer method of work was more effective than the older is indicated by the fact that the average time required to make an assembly under the older method was 0.0838 min., whereas the average time for the newer method was 0.0546 min., a saving of 35 per cent. However, it is obviously difficult to state with any degree of certainty whether the saving was due to the fact that both hands were employed in productive work at the same time or whether the new arrangement of the bins or some other change was the determining factor. Investigations such as this therefore cannot be considered satisfactory evidence that simultaneous bimanual movements are more effective than independent single-handed movements.

What is needed is the study of a task wherein exactly the same movements can be made either with the two hands simultaneously or with each hand separately in time. Barnes, Mundel, and MacKenzie have provided some pertinent evidence.[6] These investigators had subjects pick up screw nuts and drop them in a receiver hole. In one part of the investigation the nuts were picked up one at a time by one hand and in another part they were picked up two at a time with simultaneous movements of both hands. Two different types of container bins for the nuts were employed. One was a simple box-shaped bin, and the other a bin equipped with a tray that permitted the nut to be slid to the edge for greater ease in grasping. The results are given in Table 11-5. It is appar-

Table 11-5. The Times Required to Perform the Simple Operation of Picking up Screw Nuts from a Bin and Dropping Them into a Hole When Done as a Single-handed and as a Bimanual Operation

Type of operation	Box bin			Bin with tray		
	Time to grasp nut, sec.	Time to complete cycle of work, sec.		Time to grasp nut, sec.	Time to complete cycle of work, sec.	
		1 nut	2 nuts		1 nut	2 nuts
Single-handed operation:						
Right hand	0.481	1.138	2.276	0.294	0.858	1.716
Left hand	0.526	1.191	2.382	0.289	0.936	1.926
Simultaneous bimanual operation:						
Right hand	0.792		2.958	0.432		2.226
Left hand	0.719			0.410		

ent from the findings that, regardless of the type of bin employed, the single-handed operation was quicker than the bimanual operation. This investigation throws doubt on the conclusion that simultaneous motions of both hands are generally more effective than independent movements of each hand.

Curved Motions versus Straight-line Movements. Barnes found that in a simple act in which the individual carried an object in his hand away from the body, then stopped, changed direction, and carried it back toward the body, 15 to 24 per cent of the time was consumed simply in changing direction.[1] If the movements of the hand could be arranged so that no stop was necessary, as by curved motions, this waste in time could be eliminated. Barnes has shown how the application of this principle effected a marked saving in time.[2] In folding sheets of paper in one plant the worker folded the sheet over with her left hand and, with a creasing tool in her right hand, she struck the fold in the middle and creased it. She moved the tool away from her body along the fold to the further edge of the paper; then reversing the direction of movement of her hand, she moved the tool back along the fold toward her body. In an improved procedure the worker folded over the sheet of paper with the left hand in the previous manner, but she struck the fold with the tool on the near edge and moved the tool away from herself in an S-shaped curve, so that at the end of a curved movement the hand was brought back toward the body. The old method required 0.009 min. to make a fold, whereas the new method required only 0.005 min., a saving in time of 44 per cent.

Farmer found that curved motions were a more effective method of work in candymaking.[13] In the original method of work, candy centers were picked up from a tray, dipped into a pot of melted sugar, and covered with sugar by means of a series of short zigzag motions. Under an improved method of work the motions of carrying the candy center to the pot of melted sugar and moving it through the syrup were combined into a continuous curved movement. By this means the speed of production was increased by 27 per cent. The older workers, however, had used the original method for a number of years, and many of them were reluctant to change to the new method. With new employees the new method was markedly effective, and after a training period they were producing 88 per cent more than were comparable workers using the old method. It is apparent from this investigation that, even though a new method of work may be demonstrably superior to an old one, it may be difficult to establish it owing to the conflict created with well-established habits.

Symmetrical versus Asymmetrical Movements. The position generally held in time-and-motion economy is that, when simultaneous movements

of both arms are employed, the most effective procedure is to have the arms move in opposite and symmetrical directions. According to this principle, if the worker is to reach into bins simultaneously with both hands in order to pick up parts for an assembly operation, it would be more effective to have the bins placed one to the right of the worker and the other to the left than to have both bins placed on either the right or the left side. Evidence cited for the validity of this notion is usually obtained from investigations in which there are differences between the two methods of work other than merely the symmetry of the motions. As yet no valid conclusions can be drawn concerning the effectiveness of symmetrical as compared with asymmetrical motions. If coordination of the eyes and hands is required, as is likely to occur when the movements are made in picking up objects, it would appear that symmetrical movements in opposite directions probably would be less effective than movements, whether symmetrical or not, permitting operation of the hands in contiguous positions. Further investigations with regard to the relative effectiveness of symmetrical and asymmetrical movements need to be conducted for a variety of different types of tasks before satisfactory conclusions can be drawn.

Barnes and Mundel studied the effectiveness of simultaneous movements of both hands in relation to the direction of the movement.[5] In one of their investigations, subjects moved their hands away from and toward the body, the hands moving to specified points on the workbench as in reaching for objects in bins. The results, shown in Table 11-6, indicate that for the straight-line movements investigated, the direction of the movement was of little or no importance in so far as speed of movement was concerned.

Table 11-6. The Effectiveness of Simultaneous Straight-line Movements of Both Hands in Certain Opposite and Symmetrical Directions

Angle between path of movement and plane parallel to chest, deg.	Relative time to complete 1 cycle of work	Relative error of movement
90	100	100
60	95	139
30	97	189
0	104	210

Grasping Materials. The act of grasping or picking up small objects is characteristic of a wide variety of jobs. In clerical work there are papers, folders, clips, etc., to be picked up or held. In the mechanical occupations fixtures, belts, washers, rivets, etc., are grasped and picked up preparatory to such acts as assembly and insertion. Owing to the fact that this simple operation is involved in so many different jobs it warrants special consideration.

The different ways in which an object can be picked up will vary with the nature, size, and structure of that object. The operation of grasping cannot be thought of as one that can be performed in only one or two simple ways. It is beyond the scope of the present discussion to consider and evaluate all of the different possible ways of grasping materials. Only the results of one investigation will be presented here to indicate the sort of findings that may be obtained from a systematic study.

Barnes and Mundel compared two different methods of grasping and picking up small objects.[4] In their investigation the subjects picked up washers ranging in thickness from $\frac{1}{32}$ to $\frac{1}{2}$ in., the time required for grasping being noted. The methods of grasping evaluated were the pinch grasp, wherein the forefinger was placed on the edge of one side of the washer and the thumb on the edge of the opposite side; and the hook grasp, wherein the forefinger was placed on top of the washer and the thumb slightly under it. The results showed that the relative superiority of the methods is a function of the size of the object. The speed with which the washers could be picked up when the hook grasp was employed did not vary significantly with their thickness. The pinch grasp, however, was found to be much slower with thin washers and faster with thick ones.

EFFECTIVE ARRANGEMENT OF THE WORKPLACE

A number of principles for optimal arrangements of the workplace have been suggested, and in many instances significant savings have been effected by their application. To a large extent these principles are rule-of-thumb procedures and as such have not been subjected to scrutiny by scientific methods. In some cases the principles are simply statements of the obvious or what would be obvious to anyone subjecting a particular workplace to systematic examination. It is in this sense that time-and-motion economy has been helpful in developing productive arrangements of workplaces. Rather than summarizing here the rule-of-thumb principles of arrangement, the results of several investigations that are typical of the scientific approach to pertinent problems of the workplace will be described.

Pre-positioning of Tools. By pre-positioning is meant the placing of a tool in such a position—as by means of a special holder—that it will be grasped in the manner in which it is to be used. The familiar fountain-pen desk set is an example of pre-positioning. The pen is slanted at the angle at which it is normally held in the hand for writing. Similarly on assembly jobs where a screw driver is constantly used it may be suspended just above the work by a coil spring attached to the upper part of the handle. With such an arrangement the worker does not have to search for the tool or even look up from the task in hand but needs simply to reach up, grasp the handle, which is in a constant and familiar position, and bring it down.

The effectiveness of pre-positioning tools has been demonstrated by Barnes and Mundel.[4] In one of their experiments a spiral-ratchet screw driver was placed in the workplace in three different degrees of pre-positioning. In complete pre-positioning it was held over the workplace by a coil spring. In partial pre-positioning it was held by a fixture on the bench to one side of the working space but in a position permitting ease of grasping in the manner in which the tool was used. Finally the screw driver was simply placed on the workbench without pre-positioning. The effectiveness of these three arrangements was measured in terms of the speed with which bolts could be tightened by means of the screw driver. The number of bolts tightened per minute when the tool was completely pre-positioned was 24.0, when partially pre-positioned 19.5, and when not pre-positioned 16.4. These findings clearly demonstrate the effectiveness of pre-positioning.

Types of Bins. So many industrial jobs require the worker to pick up objects from containers preparatory to utilizing them in some way that it will be worthwhile to consider the effectiveness of different types of bins. Barnes and Mundel evaluated three different types.[4] One was a simple box-shaped bin with a flat bottom that required the worker to

Table 11-7. The Time Required to Select and Grasp a Small Object from Different Types of Bins and Drop It into a Hole

	Box bin		Hopper-type bin		Bin with tray	
	Nuts	Screws	Nuts	Screws	Nuts	Screws
Time to grasp object, sec.	0.392	0.321	0.377	0.352	0.234	0.279
Time for 1 cycle of work, sec.	0.888	0.968	0.826	0.940	0.696	0.857

dip his fingers into the contents in order to pick up an object. The second type was the hopper-fed bin which is similar to the box bin except that the bottom is rounded and there is a constant feeding of materials so that searching for an object is presumably reduced. The third type of bin was similar to the hopper bin, but with the added feature of a lip tray on the near side so that the fingers could slide an object out to the edge of the tray for ease in grasping. The results given in Table 11-7 indicate that, in so far as time to grasp an object is concerned, the box bin and the hopper-fed bin show no consistent differences, whereas the bin with a tray is superior to both other types. Similar results were obtained when the times for total cycles of work were considered.

The Working Area. Discussions of effective arrangements of the workplace invariably stress what is known as the maximum working area and warn against the placement of bins, tools, and machine controls outside of this area. If a person is seated in an erect position at a workbench and describes arcs on the table top with his extended arms, the circumscribed area is called the maximum working area. Should the worker be required to reach outside this area, he must change his posture by bending his trunk either forward or to one side. To hold that all manual activity should be performed within these confines presupposes that any change in body posture is undesirable. However, a case can be made for just the opposite point of view. It can be argued that the use of the large trunk-muscle groups in bending the body, as in reaching, will effectively relieve the shoulder and arm muscles of some of their work. Over a long period of time the occasional use of trunk muscles in conjunction with arm muscles, as compared with the use of only arm muscles, might give as great or even greater productivity. In view of inadequate data this question cannot be answered at the present time. Certainly a set of definitive experiments are needed before it can be definitely concluded that all work should be performed within the so-called maximum working area.

As might be expected, variation in the location of tools and materials within the maximum working area produces changes in the rate of work. Barnes investigated the speed of movement of the arm away from and toward the body in relation to the distance of movement, all distances being well within the maximum working area.[1] As will be seen in Table 11-8, the time to complete a movement increased as the length of the movement increased. It is important to note, however, that the time required for the longer movement was not increased in proportion to the increase in length of the movement. Rather it took relatively less time to make longer movements than shorter movements. In other words, the velocity or speed of movement was greater for longer than for shorter movements. The results also show that the percentage of time required

Table 11-8. A Comparison of Long and Short Movements of the Hand and Arm

Distance, in.	Velocity		% distance required for acceleration of movement		% distance required for deceleration of movement	
	Toward body	Away from body	Toward body	Away from body	Toward body	Away from body
5	12	11	35	34	35	36
10	20	20	35	37	37	35
15	25	26	34	39	39	33

for acceleration and deceleration of the movements varied systematically with the distance traversed.

In another part of the experiment the subject reached for a part that he carried back to the work space for insertion into a fixture. In this experiment the distances investigated were up to 24 in., which is practically the outer limit of the working area. As shown in Table 11-9, the

Table 11-9. A Comparison of Long and Short Movements of the Hand in an Industrial Operation

Distance between bin and fixture, in.	Movement of empty hand from fixture to bin		Carrying part from bin to fixture	
	Time, sec.	Velocity	Time, sec.	Velocity
8	0.222	36	0.270	30
16	0.264	61	0.384	42
24	0.348	69	0.420	57

results obtained were similar to those of the first experiment; *i.e.*, increasing the length of the movement required longer time for the movement. It will be seen, however, that there was an increased speed of movement as indicated by the velocity.

PSYCHOLOGICAL FACTORS IN EFFECTIVE DESIGN OF TOOLS AND EQUIPMENT

The designing of tools and equipment has largely fallen to the engineer. To some extent this has been due to a lack of interest on the part of psychologists, but to a large extent it has been due to the fact that the design of tools and equipment has been dictated by mechanical and financial considerations rather than by human factors. There is no machine so automatic and so perfect that no human control or direction is required.[9] Indeed there is something to be said for the point of view that tools and machines are simply mechanical extensions of the human organism, created for the purpose of making more rapid, more accurate, or stronger responses. Viewed in this light it is apparent that the effectiveness of industrial equipment is a direct function of the degree to which human factors are taken into account. The mechanical and financial aspects, of course, must be considered, but it is also true that various savings and benefits will be obtained by designing the tools and equipment in terms of the characteristics of the individual who will use them in industrial work.

It cannot be assumed that merely because a particular type of equipment permits of greater mechanical advantage or speed for the operator, it will necessarily result in greater effectiveness of work than other equipment that is less adequate mechanically. Equipment does not operate by itself, but to some degree at least is controlled by an operator. This means that other than purely mechanical factors need to be considered in the design of equipment. In the following paragraphs design will be considered in relation to human capacities and limitations, the minimizing of human error, and acceptability by the worker.

Design in Relation to Human Capacities and Limitations. Persons whose jobs require them to use tools or machines are perpetually amazed at their inadequacies in so far as human capacities are concerned. The damnations directed by workers at their tools and machines can be taken as signs of more than the workers' personal deficiencies. Such verbalization may be a sign that the equipment is inadequate. Consider how many machines have all the controls on the left-hand side, and how many wrenches have handles so short that even a Sandow could not cinch a bolt with them.

The various motor and perceptual capacities of human beings are the basic data that should be considered in the designing of any equipment. They are the factors that should enter into the determination of the placement, form, and mode of operation of controls, dials, indicators, and other

parts of equipment with which the worker is concerned. If under ordinary circumstances a typical worker in pulling a lever does so with a force of about 70 lb., then to pull the lever on his machine should not require a force of more than 70 lb. If a worker cannot read figures ⅛ in. high on a dial 10 ft. away from him, then the figures should be made larger. Normative data with respect to various motor and perceptual capacities of the worker provide guides for the designer of tools and equipment.[8] Knowledge of the range of individual differences in strength of pull, reaction time to auditory stimuli, visual acuity, etc., is important. When such information is not available, specific investigations need to be undertaken to collect it.

The common typewriter is an instance of a most flagrant violation of the principle of adapting the machine to the worker's capacities. On the basis of tests of tapping and analyses of the errors made by the different fingers, the ability of each finger for typing was estimated.[20] In the top row of Table 11-10 the relative amount of work each finger and hand

Table 11-10. A Comparison of the Load Carried by Each Finger and Hand on the Ideal and on the Present Typewriter Keyboards

Kind of load	Left-hand load, %					Right-hand load, %					Total both hands, %
	Finger				Total	Finger				Total	
	4	3	2	1		1	2	3	4		
Ideal load	10.8	11.4	12.3	13.0	47.5	13.9	13.9	12.5	12.2	52.5	100
Present load	10.2	8.3	18.9	19.4	56.8	18.8	8.1	12.6	3.7	43.2	100
Ratio of present to ideal	94	73	154	149	120	135	58	101	30	82	

should do—the ideal load—is given. These figures may be contrasted with those in the second row giving the load required by the arrangement of the keys on the present typewriter. From this table it can clearly be seen that the present typewriter is inadequate in the distribution of work it assigns to each finger and hand. For example, the fourth finger of the right hand only does one-third of the work that it should, while the second finger of the left hand does over 50 per cent more work than it should. The left hand does one-fifth more work than it should, and the right one-fifth less. On the basis of studies such as these, keyboards yielding maximal mechanical advantage have been constructed. Tests made on one such typewriter revealed its advantage in terms of reduc-

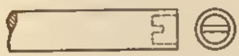
tion in learning time, increase in speed of typing, reduction in errors, and reduction in work decrement.[11]

Design to Minimize Probability of Human Error. In discussions of highway design with regard to safety problems it has been said that the most effective approach is to develop a highway whereon the likelihood of the automobile driver getting into difficulties is at a minimum. In other words a highway should be so constructed as to reduce all potential sources of friction in the flow of traffic. Human error is reduced by eliminating those situations where the individual has an opportunity of making an error. This point of view can be carried directly over to the field of design of tools and equipment. The performance of an individual is never perfectly consistent. When possible, inaccurate responses should be detected and described, and steps taken to obviate the possibility of their occurrence.

Accidents, of course, are the most striking results of human error in industry. In a later chapter concerned with accidents it will be pointed out that the analysis of workers' errors is an important step in the development of adequate safety measures. Errors, however, may lead to a wide variety of other types of detrimental results. Errors may bring about poorer quality of production or reduced speed of production. A misread gauge may cause damage to equipment or result in a faulty product. Too strong pressure on the control lever of a drill press may cause the bit to break or the hole to be drilled out of round. Again, the interest is in the various motor and perceptual responses of the individual, but here the concern is with their possible inadequacies or inaccuracies. If tools and machines are designed so that inadequate motor and perceptual responses are minimized, then work involving these tools and machines will be more effective.

Tightening screws with a screw driver is a common example of how human error leads to ineffective performance. In tightening a screw the individual should keep the tool in direct line with the screw. If the operator tilts the screw driver out of this direct line, the bit is apt to slip from the slot in the screw, resulting in possible damage to fingers and materials and slower production, since the bit must again be set in the slot to complete the task. It seems clear that a screw driver designed to minimize or eliminate this slippage would be superior to the ordinary type of tool. Barnes and Mundel evaluated two different screw drivers designed to minimize slipping from the screw head, and compared them with the ordinary screw driver.[4] The types of bits they studied and their findings are given in Table 11-11. From these results it can be seen that the use of screw drivers that reduce the slipping of the bit from the screw head results in higher levels of production. Although the ef-

Table 11-11. The Effectiveness of Screw Drivers with Different Types of Bits

Type of bit	Number of screws tightened per min.	
	With screw already inserted	With screw started by operator
	10.7	5.7
	11.5	6.3
	12.1	6.6

fectiveness of these bits in other respects was not investigated, it is likely that they also would reduce damage to fingers and to materials.

Design to Achieve Workers' Acceptance of Tools and Equipment. One problem with respect to the design of tools and equipment that has received little attention is the extent to which they are acceptable to the worker. The safety engineer engages in a never-ending struggle to persuade workers to use goggles, gloves, and other similar equipment. The chemical worker who is provided with devices to help him gauge the effectiveness of his mix will ignore them in preference to a taste test. The mechanic who is furnished with a special set of wrenches will nevertheless use a pair of ordinary pliers.

Obviously there are many different factors that determine whether or not workers will accept a particular piece of equipment. They may lack confidence in it, either through unfamiliarity or because of lack of experience. Some equipment is not comfortable and some may actually limit activity. In certain instances a piece of equipment may conflict with custom or be otherwise socially unacceptable. Thus welders' masks are made acceptable to women workers by painting them with colorful and attractive designs. Whatever the reason for lack of acceptance, it is apparent that no matter how mechanically effective or desirable a piece of equipment is, it is of no value if it is not accepted and used by the worker.

PROBLEM AREAS IN THE DESIGN OF EQUIPMENT

Problems in the development of equipment fall into three areas.[15] One set of problems, largely of a perceptual sort, revolves around information-giving systems such as dials and warning signals. A second set of problems arises out of the design of control systems such as handles, levers, and hand wheels. A final set of problems, concerned with organizational systems, bears upon the relationships among groups of men and machines. These three problem areas are by no means independent. For example, the manual controls of an airplane for rudder and ailerons are not only control systems but also information-giving systems since the amount of resistance experienced in adjusting them furnishes the pilot with cues about the position of the airplane and the forces operating upon it.

Information-giving Systems. Information can be conveyed to an individual through any of the various sense modalities. However, in most cases information-giving systems employ either visual or auditory presentations. When designing an information-giving system the basic question to be asked is: What kind of information does the operator really need? Usually little thought is given to this problem, and as a consequence the operator either gets more information than he can use or not enough. The operator may need a precise quantitative reading in order to find out exactly what a particular state of affairs is. Thus the driver of an automobile wants to know just how fast his vehicle is traveling. The judgment as to whether the particular speed is too fast will depend upon the legal speed limit, the traffic conditions, the time available to reach a particular destination, etc. Or, as a basis for his judgment the operator may only need to know that a particular state of affairs deviates from a standard or desired state of affairs in a given direction and in a given amount. Thus in maintaining pressure at a given level in a boiler, the engineer would need to judge when to release steam to reduce pressure that is too high, or when to increase the fuel to raise pressure that is too low. Finally, the operator simply may need to know whether a particular state of affairs exists or does not exist. Warning signals, such as the horn which tells the pilot during the landing maneuver that his wheels are still retracted, are examples of such "go–no go" systems. There are so many different kinds of information-giving systems that it will be impossible to review them all here. As a consequence one typical problem will be considered.

One of the most commonly employed types of information-giving systems is the visual display, *e.g.*, dials. Sleight studied the various types of common dial faces shown in Fig. 11-1.[23] Effectiveness of the dials was judged in terms of the errors made in reading them during short ex-

posures. The percentage of readings that were in error for the five types of dials tested were about as follows: open-window 1 per cent, round 10 per cent, semicircular 17 per cent, horizontal 28 per cent, and vertical 36 per cent. The open-window dial clearly is the best in terms of accuracy in reading. As might be expected errors increased when the indi-

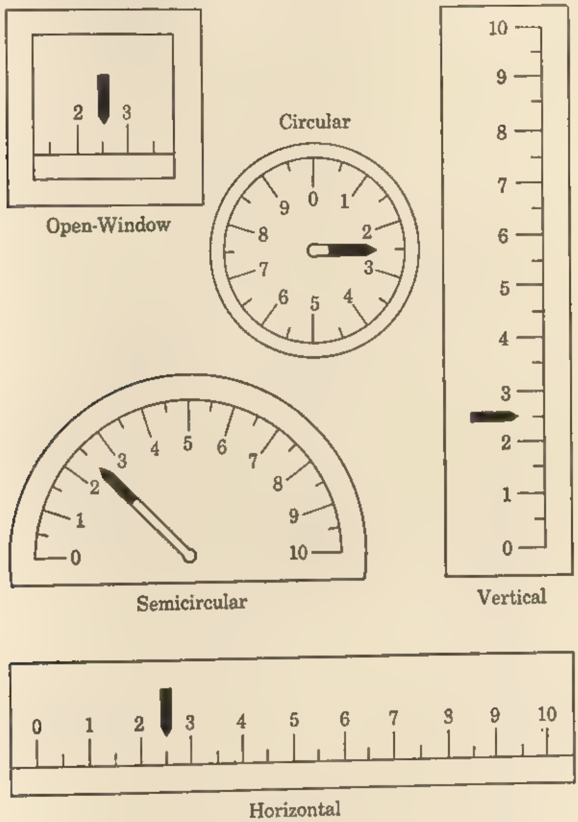


Fig. 11-1. Examples of common types of dials.

vidual had to interpolate values between the marked numbers. With the horizontal and vertical dials more errors occurred when the pointer was toward either of the ends of the scale than when it was in the middle range. It is to be noted that when information other than that of amount is to be obtained from dials, the order of effectiveness found by Sleight might not hold. For example, with direction indicators such as compasses, the open-window type of dial might be more confusing than a circular dial.[15]

Control Systems. An analysis of the situations in which control systems are employed indicates that they are utilized to effect five different types

of reactions.[8] One type is a static reaction, wherein something is maintained in a fixed position. The holding of a nail that is being hammered is one example of a static reaction, and the holding of a lathe tool against turning material is another. Poor static reactions show up in the form of tremor or shaking. Providing support minimizes tremor.

A second type of reaction is a blind-position reaction. In such reactions the individual does not know the accuracy of his response until it has been accomplished. These reactions are of particular importance in the location of machine controls. For example, as the pilot is landing an airplane he is unable to look for the lever which lowers the flaps but must reach for it blindly. Studies of blind reaching indicate that controls directly in front of the individual are most accurately located.[14] For other than frontal positions there is less accuracy in locating controls above the shoulders than below them. In reaching for controls placed in either high or low positions the individual usually does not reach far enough.

A third type of reaction involves visual-positioning movements, wherein the individual gets information visually as he manipulates a control. Setting the hands of a watch to the correct time of day is an illustration of this type of reaction. As the stem of the watch is turned the position of the hands changes with respect to the markings on the face of the watch, and the individual knows whether he has turned the stem too far or not far enough. Analysis of visual-positioning movements indicates that they have three components, namely, reaction time (the time between the starting signal and the beginning of the movement), primary-movement time (the time between beginning the movement and grossly approaching the position desired), and secondary-movement time (the time taken to make final minor corrective movements).[8]

Repetitive movements such as turning a crank are a fourth type of reaction. Studies of repetitive movements of controls show that the effectiveness of operation is not determined by a single factor, but rather that speed of movement, direction of movement, resistance, inertia, and size, position, and height of the control are all determining factors.[2,19]

A final type of reaction, continuous-adjustive movements, may be thought of as a special case of visual-positioning movements. In the latter a given and set mark is sought, while in the former the mark changes and the individual must continuously adjust the controls to it. Steering an automobile is an illustration of a continuous-adjustive movement since the stimulus, the road, and traffic conditions are continually changing. The driver continuously adjusts the steering wheel to maintain the progress of the vehicle within the boundaries of the roadway.

Important Characteristics of Controls. Three characteristics of controls that deserve special attention are the force with which they can be operated, their distinguishability, and their realism. Studies of the design

of screw-driver handles can be cited to illustrate problems with respect to force of movement. When the effectiveness of screw drivers is measured in terms of the force with which the tool can be turned, it appears that handles with uniform diameter throughout their length are about 20 per cent more effective than those in which the diameter of the lower part is less than that of the upper part.[2] With light work, however, it is desirable to have the lower part of the handle smaller in diameter than the upper part.[21] Such a shape permits rapid insertion of the screw by rotating the lower part of the screw-driver handle between the thumb and the first-finger and permits a final tightening by use of the increased leverage provided by the large diameter of the top portion.

Many machines require the operation of a number of different controls. The control system of an airplane is an excellent example. When the operation requires the individual to move rapidly among the controls, their distinguishability becomes particularly important. Controls can be distinguished from one another by their location, color, size, and shape. The problem resolves itself into one of designing the controls so they can be readily distinguished by the operator. In spite of the fact that a wealth of evidence has been accumulated on this problem, little of it has been employed.[16]

Realism in controls revolves around the problem of naturalness.[8] If a control knob is supposed to move a dial pointed in a clockwise direction then it would appear obvious that the knob should also move in a clockwise direction. Similarly it would appear desirable to have the lever which controls a lifting machine moved up by the operator when the lift is to be raised and moved down when the lift is to be lowered. Realism as a principle in the design of control systems would seem to be so important that it would never be ignored, yet many tools and machines require movements that are quite unnatural. Furthermore, in many cases the natural movement is by no means obvious. For example, suppose that the pointer on a vertical dial is controlled by a lever that is either pulled toward the body or pushed away from it. In order to move the pointer up is it more natural to pull the lever or to push it? Research conducted in the area of realism in the operation of controls has yielded many unexpected results. However, the field is so new that as yet no useful set of principles has been established.

Organizational Systems. The foregoing discussion of matters connected with design has dealt with the problems quite microscopically. The concern was with individual pieces of equipment and the specific responses of the worker in operating this equipment. A more macroscopic view produces problems of greater complexity involving interactions among men and machines, men and men, and machines and machines. The integration of the activities of an organization comprising men and equip-

ment involves more than merely the selection and placement of personnel and the design of equipment. Problems connected with organizational systems frequently are difficult of solution, and in many instances the conceptualization is beyond human capacity, at least in the present state of ignorance.

Perhaps the simplest problem in organizational systems involves the layout of the workplace. Here the concern is with placing tools and materials in relation to the worker so that his work is most effective. More complex problems arise from the interrelations among the activities of a number of men and machines. In the terms of the industrial engineer, these problems involve what he calls process analysis.[2] The concern may be with the spatial arrangements of men and machines, with placements being made to effect the quickest flow of work from one man or machine to another. The most complicated problems are confronted in structuring the functions and activities of a large organization consisting of many men and machines. The fire-control system on a battleship is one example, and the operation of a radio-controlled-taxicab system is another.

In situations where the process or activity is standardized the problem is relatively simple. The arrangement of the workplace for a routine assembly job will take into consideration many of the matters considered earlier in this chapter such as normal limits of reach, type of bin, etc. In an assembly line operation where subassemblies come to a main line the timing of the operation must be so planned that the subassemblies will reach the main line at the proper rate and time. Situations such as these can be described by means of motion-cycle charts, process charts, or flow charts.

With unstandardized activities the problems are more difficult. A motion-cycle chart is of little use in describing the activities of the crew of an airplane in relation to their equipment. Here there is no routine. Depending upon circumstances and the course of events, the various members of the crew refer to different gauges, operate different controls, and give different kinds of instructions and information to each other. In order to describe the interactions in such situations, some investigators have suggested that linkages be analyzed.[8] Whenever one individual, whether man or machine, interacts with another, a linkage is said to occur. Linkages can be analyzed in terms of frequency, duration, and importance, and a chart can be constructed to depict the integration of the activities involved.

In organizational systems it is not enough just to know the characteristics and capacities of the persons and machines that comprise it, but in addition these characteristics and capacities must be evaluated in terms of the total functioning of the organization. For example, it

might be possible to design a machine that takes over the functioning of a human being, provided that the service personnel available are capable of maintaining the machine in adequate repair. In the long run, however, it might be more effective to have a slower and simpler machine that can be maintained by its operator rather than one that requires a maintenance specialist.

REPETITIVE WORK

The last three decades have witnessed a marked shift in industry from the more complex, varied, and individually planned work of the skilled craftsman to the simpler, uniform, and repetitive work of the semiskilled production-line operator. Particularly during the last war, many jobs that hitherto had been performed by skilled workers were reanalyzed and divided into several different subjobs, each of which could be performed by persons of far less training and experience. The reasons offered for this change have been many, but economic and technological considerations appear to have been the major factors. In so far as the human side of the picture is concerned, the only justification offered has been the reduced amount of training required by the worker in order to perform the job adequately. For an industrial task requiring a high degree of skill the training time will be in terms of years, whereas for a simpler task the training time will be in terms of hours, days, or at the most, weeks. From the psychological point of view this is clearly an inadequate justification for the substitution of the simpler repetitive tasks. Factors such as the morale, adjustment, and welfare of the worker should also be considered.

The increasing number of simple repetitive tasks in industry has impelled many writers to fill page after page with explications of the evil effects of the system. Unfortunately most of these discussions are of a romantic sort and contribute little toward a scientific understanding of the basic problem. It should be pointed out that repetitive work is not an innovation of the industrial age. The pneumatic-hammer operator and the multiple-loom operator of today are probably no more or no less bored by their work than were the ancient hewer of stone and the weaver of cloth.

Repetition and Monotony. Implicit in criticisms of repetitive work is the assumption that all work of this kind is necessarily monotonous, and monotonous for all workers. It cannot be denied that repetition is a factor in monotony. In an automobile factory 67 per cent of men engaged in single operation jobs reported their work as boring, 56 per cent of those whose jobs involved two to four operations, and only 30 per cent of those whose jobs involved five or more operations.[27] However, the evidence certainly does not indicate that repetition, in and of itself, is the major determinant of monotony in industrial work. The facts in oppo-

sition to this thesis are first, that for the same repetitive task all individuals are not bored; secondly, that changing the conditions of work (*e.g.*, permitting talking), but not the repetitive nature of the task, reduces or eliminates monotony; thirdly, that changing the nature of the incentive (*e.g.*, paying on a piece-rate rather than a time-rate system), but not the task, reduces or eliminates monotony; and, fourthly, that on varied and nonrepetitive tasks some workers report that they are bored.[28] This is not to argue that repetition is not a factor leading to monotony, but rather that the unvarying nature of the task leads to monotony only with a particular individual, under certain conditions of work, and with an inadequate motivating situation. The notion that there is an inherent dislike for repetitive work has never been substantiated and should no longer be maintained.

Degree of Automatism of the Work. Industrial tasks can be considered in terms of the degree to which they are demanding upon the worker's attention. Although obviously all tasks require some attention, for some tasks, either because of the simplicity of the operations involved or because of their habitual nature, the worker needs to devote very little attention to the job. Certain assembly and packing operations are of this sort. Such jobs may be performed almost completely automatically. On the other hand there are certain tasks that, at least for extended periods of time, require almost the entire attention of the worker. Thus the bus operator on a busy run and the clerk who adds and balances a series of figures must necessarily concentrate closely on the task in hand. Activities such as these cannot possibly be performed in an automatic fashion.

In between these two extremes are jobs requiring intermittent close applications of the worker's attention. It may be that for a given job certain of the activities can be done in an automatic fashion whereas others present novel problems, or perhaps the job operations are of such a nature that they can usually be performed in an automatic fashion with frequent close attention when checks are made of the quality of the work. Such work can be called semiautomatic since it is partly attention demanding and partly automatic. An illustration of work of this kind is that of the bottling-machine operator. With certain types of machines the bottles are placed into the machine by hand, and the amount of fluid delivered into the container is controlled by the worker. These operations can be done automatically by him and require little attention. However, frequently the supply of bottles is interrupted, or he receives an imperfect bottle, or he misjudges the amount of fluid delivered. These variant situations occur at irregular and unexpected times and require an immediate and fully conscious adjustment.

Semiautomatic jobs are found at all occupational levels and are not restricted to the manual occupations. Consider the job of the airplane pi-

lot. To be sure, while flying a traffic pattern and in landings and take-offs, his entire attention to the task in hand is required. The amount of time consumed by such activities, however, is only a relatively small proportion of his total working time. While flying cross country on a set course and with at least reasonably good weather conditions, he needs only occasionally to check closely his instruments and surroundings. Under these conditions his work is truly semiautomatic.

Degree of Automatism and Monotony. Contrary to expectation it is not the job that can be done in a completely automatic fashion that is most likely to produce monotony. Rather, boredom is most likely to occur when the individual finds it necessary to devote some but not all of his attention to his work. When the individual can perform his work with little or no attention to it, or when the work demands his complete attention and interest, the job is less likely to be reported as being monotonous. In other words monotony is less likely to occur with work that is completely automatic or with work that demands a high degree of attention, and more likely to occur in the case of semiautomatic work.[10, 28]

The explanation of this state of affairs is not difficult to understand. On the one hand, a job that is completely automatic leaves the worker free for reverie or, if the working situation permits, conversation and social intercourse with his fellow workers. On the other hand, work that requires complete attention does not allow for reverie, and unless it has a high degree of intrinsic interest for the individual it will not be continued without letup for an extended time. Compared with these, semiautomatic work neither permits freedom of thought nor is sufficiently demanding to keep the individual's mind fully occupied.[28] Because of its very nature semiautomatic activity may well have low interest value, but because it does on occasion demand a high degree of attention it will not permit the worker to make the wholly satisfactory attentive adjustment that he can make to highly automatic work.

It must be understood that the relationship described above between the degree of automatism of the work and monotony is only a general tendency. Some tasks that are either completely automatic or require a high degree of attention are reported as boring. In the one case the worker may not succeed in making a satisfactory adjustment in terms of extra-job activities such as conversation with fellow workers, and in the other, even though the work receives the greater part of the worker's attention, it still may be intrinsically uninteresting.

REFERENCES

1. Barnes, R. M.: An investigation of some hand motions used in factory work, *Univ. Iowa Studies Eng.*, No. 6, 1936.

2. Barnes, R. M.: "Motion and Time Study," Wiley, 1946.
3. Barnes, R. M., and M. E. Mundel: Studies of hand motions and rhythm appearing in factory work, *Univ. Iowa Studies Eng.*, No. 12, 1938.
4. Barnes, R. M., and M. E. Mundel: A study of hand motions used in small assembly work, *Univ. Iowa Studies Eng.*, No. 16, 1939.
5. Barnes, R. M., and M. E. Mundel: A study of simultaneous symmetrical hand motions, *Univ. Iowa Studies Eng.*, No. 17, 1939.
6. Barnes, R. M., M. E. Mundel, and J. M. MacKenzie: Studies of one hand and two hand work, *Univ. Iowa Studies Eng.*, No. 21, 1940.
7. Beeby, C. E.: An experimental investigation into the simultaneous constituents of an act of skill, *Brit. J. Psychol.*, **20**, 336, 1930.
8. Chapanis, A., W. R. Garner, and C. T. Morgan: "Applied Experimental Psychology," Wiley, 1949.
9. Coakley, J. D.: Human operators and automatic machines, *Personnel Psychol.*, **3**, 401–411, 1950.
10. Davies, A. H.: The physical and mental effects of monotony in modern industry, *Brit. Med. J.*, No. 3427, 472–476, 1926.
11. Davis, D. D.: An evaluation of the simplified typewriter keyboard, *J. Bus. Educ.*, **11**, 2, 1935.
12. Fairchild, M.: Skill and specialization, *Personnel J.*, **9**, 20–71, 128–175, 1930.
13. Farmer, E.: Time and motion study, *Ind. Fatigue Research Bd.*, No. 14, 1923.
14. Fitts, P. M.: Psychology and aircraft design, *Mech. Eng.*, Feb., 1947, 135–141, 163.
15. Fitts, P. M.: "Psychological Research on Equipment Design," Army Air Forces Aviation Psychology Program Research Reports, No. 19, 1947.
16. Gilbreth, F. B.: "Motion Study," Van Nostrand, 1911.
17. Gilbreth, F. B.: "Applied Motion Study," Sturgis & Walton, 1917.
18. Hartson, L. D.: Analysis of skilled movements, *Personnel J.*, **11**, 28–43, 1932.
19. Helson, H.: Design of equipment and optimal human operation, *Am. J. Psychol.*, **62**, 473–479, 1949.
20. Hoke, R. E.: Improvement of speed and accuracy in typing, *Johns Hopkins Studies Educ.*, **7**, 1–42, 1922.
21. Hunt, L. I.: A study of screwdrivers for small assembly work, *Human Factor*, **8**, 70–73, 1934.
22. Jenkins, J. G.: "Psychology in Business and Industry," Wiley, 1935.
23. Sleight, R. B.: The effect of instrument dial shape upon legibility, *J. Appl. Psychol.*, **32**, 170–188, 1948.
24. Stetson, R. H., and J. A. McDill: Mechanisms of the different types of movements, *Psychol. Monograph*, No. 145, 1923.
25. Taylor, F. W.: "Common Sense Applied to Motion and Time Study," Harper, 1911.
26. Tolman, E. C.: "Purposive Behavior in Animals and Men," Century, 1932.
27. Walker, C. R., and R. H. Guest: "The Man on the Assembly Line," Harvard University Press, 1952.
28. Wyatt, S., J. A. Fraser, and F. G. L. Stock: The effects of monotony in work, *Ind. Fatigue Research Bd.*, No. 56, 1929.

CHAPTER 12

Accidents and the Safety Problem

The number of workers involved in industrial accidents is so astonishingly high that a person not familiar with the actual statistics undoubtedly would underestimate their frequency by more than one-half. The number of industrial workers killed or disabled in connection with their jobs runs well over two million per year. The injuries are estimated to cause a loss in full-time annual employment of the equivalent of about 750,000 workers, or of some 220,000,000 employee-days. To these disabling accidents must be added an untold but very large number of accidents that injure but do not disable, and still others that do not injure but result in property damage. It is apparent that just in terms of sheer frequency of occurrence the accident problem is of utmost importance. Any procedure, therefore, by which accidents can be reduced even by a fraction of one per cent must necessarily be considered of great value.

In view of the enormity of the situation, careful consideration must be given to the nature, causes, and conditions of accidents. This does not mean that the critical scientific analysis should be dropped and a popular survey analysis substituted in its place. Rather the reverse is true. It is unfortunate that in the area of accident study so many rule-of-thumb and unverified methods have been utilized. In desperation many things are considered truths that have not been subjected to scientific tests. Although perhaps no great harm is done, much good is left undone. In accident studies there must be greater awareness of the dangers inherent in inadequate data and in concepts based on opinion rather than on facts. Thinking must be clear, definitions exact, and conclusions based on sound fact, as in the study of any other problem in human behavior.

DEFINITION OF ACCIDENTS

Before factors related to accidents and their causes can be studied, it is necessary to count the number of accidents that occur under different kinds of conditions. Essential to such statistical analysis is an exact delineation of just what is to be counted. In other words the first step in

accident analysis is to define just what constitutes an accident. It will be seen that a satisfactory definition of accidents is not easy to devise, and there will be marked variations in the meaning given the term from one situation to another.

Types of Definitions. Lexigraphically, an accident is an event that takes place without foresight or expectation and results in some type of personal injury and/or damage to equipment or property. However, such a definition is too broad for use in industrial-accident analysis. The act of a saboteur in a plant fits this definition since, except to the saboteur, the act is not anticipated and results in destruction. In such cases the injury or damage would not be considered to be the result of an industrial accident. In the transportation industry the throwing of a missile by a malicious child that causes damage to a vehicle is generally considered to be an accident. But such an event really does not differ significantly from the act of an industrial saboteur. Similarly, unpredictable natural events such as heavy rains, landslides, and earthquakes will cause damages that are referred to as accidents.

A narrower definition of industrial accidents conceives them as arising directly out of the work situation, *i.e.*, from faulty equipment or the inadequate performance of an individual. In the case of the saboteur or the child, the damage is neither attributable to faulty equipment nor to inadequate performance of a worker. When natural catastrophes are the agents that produce the destructive or injurious events, there is clearly no connection with the job operations.

There are also other definitions in current use. Since a large proportion of accidents result in some claim either for personal injury or property damage, legal definitions are utilized. A further interpretation can be found in the treatments of accidents in connection with labor statistics. Here a mishap is identified as an accident only when a worker is injured and as a consequence loses time from his job.

From the foregoing statements it is apparent that different criteria are used in different cases to determine whether the mishap is to be classified as an accident or not. Confusion arises when more than one criterion is used in the same analysis and when no attempt is made to differentiate between the different criteria. The truth is that in most studies the investigator does not seem to be aware of the fact that he may be using more than one criterion.

There are three bases according to which accidents may be classified or grouped, viz., the nature of the event, the causes of the event, and the effects of the event. Generally the first or the last method of classification is employed, but, all too frequently, two or all three types of categories are used in the same analysis. Examples of these three ways of classifying accidents are given in Table 12-1. In many studies these dif-

Table 12-1. Examples of Systems for Classifying Accidents

- *Classification according to nature of the event* (trolley-car and bus accidents)
  - Traffic accidents
    - Collisions with pedestrians
    - Collisions with cars or buses
    - Collisions with motor vehicles
    - Collisions with fixed objects
    - Derail or switch occurrence without collision
  - Passenger accidents
    - Boarding
      - Boarding moving car
      - Caught or struck by doors
      - Trips, slips, stumbles
    - Alighting
      - Alighting from moving car
      - Caught or struck by doors
      - Trips, slips, stumbles
  - Occurrences on board
    - Falls, stumbles, slips
    - Miscellaneous
- *Classification according to cause of the event* (machine-shop accidents)
  - Accidents connected with machines
    - Catching of fingers, arms, clothing, etc., in machines
    - Catching of tools, guides, etc., in machines
    - Flying objects and particles
  - Nonmachine accidents
    - Falling objects
    - Objects on floor
    - Pushes, bumps, etc., by other persons, trucks, etc.
- *Classification according to effects of the event* (accidents in warehouse)
  - Damage
    - To stored material
      - Complete loss of container and contents
      - Partial damage of container and contents
      - Partial damage of container only
    - To equipment
      - Hand trucks
      - Dollies
      - Conveyor
      - Crane
  - Injury
    - Fatalities
    - Amputations
    - Foreign body in eye
    - Burns and scalds
    - Bruises, contusions, and abrasions
    - Cuts and lacerations
    - Fractures
    - Sprains and strains
    - Punctures

ferent types of classification are not distinguished. Thus in a breakdown of a given set of accidents some will be tabulated in terms of the nature of the event, others classified in terms of causes, and still others arranged according to their effects.

Need for Using Common Factors in Defining Accidents. The point should be emphasized that different accident statistics are not comparable unless the same base has been used in classifying the accidents. Comparisons of the hazardous nature of different jobs are made when the criterion used in classifying the accidents on one job is not the same as the criterion used on other jobs. Accidents in the transportation industry, where natural catastrophes play a part, are not comparable with accidents in clerical work, in which such catastrophes are never, or seldom, found. As long as investigators are unaware of this problem in classification and continue to compare accidents that are described in nonparallel categories, the true picture of the determining factors operating in accidents will be obscured.

ACCIDENT STATISTICS

One of the most difficult aspects of accident analysis is the formulation of adequate statistics. In order to ascertain which of two truck drivers is the safer operator, it clearly would be improper to compare them simply on the basis of number of accidents sustained. The accident record for one driver might cover a much longer period than that for the other. Knowing only that one driver had two accidents and the other had four accidents is not sufficient information to compare them in terms of safety in operation. It would be of some help to know that the first man had two accidents in 1 month and the second had four accidents in 3 months. Then it would be possible to equate the accident records of the two men on the basis of time and compute for each an accident rate. In this manner it would be seen that the accident rate of the first driver is 2.0 accidents per month and that of the second is 1.3 accidents per month. The latter now is seen to be the safer driver.

Determination of Accident Rates. Accident records for different individuals, conditions of work, kinds of equipment, etc., can be compared only when there is equal opportunity or exposure for all concerned. In the case of the two truck drivers mentioned above the attempt was made to equate opportunity for having accidents statistically by reducing the accidents to a monthly basis. It might be more accurate, of course, to ascertain the actual number of hours each man had driven and then compute a rate in terms of the number of hours of operation. Suppose, however, that one operator was so cautious that he drove at only half the speed of the other. Therefore, although his accident rate per *hour* of

operation might be less, in terms of *distance traveled*, or *work done*, it might be greater. It is questionable whether reducing accidents to equal times actually equates opportunity for having accidents. Equal opportunity, perhaps, should be conceived of as equal number or extent of operations, rather than equal time.

Expressing safety of performance by a rate as just described still may not equate for opportunity or exposure. The bus driver whose route covers the busy downtown district is certainly exposed to different hazards than the driver whose route runs through suburban areas. Qualitative differences in the conditions under which different workers perform the same job may introduce variations in hazards which have to be taken into account if the workers are to be accurately compared with respect to safety of performance. In a study of accidents of streetcar motormen, Kraft and Forbes found that they not only had to take into account the total distance traveled but also variations in routes, hours of the day worked, and days of the week.[26]

Variation in Results with Use of Different Types of Accident Rates. That different methods of computing accident rates may make an important difference in the conclusions drawn from an analysis is illustrated in Table 12-2. In this table are given the basic accident data and the

Table 12-2. Comparison of Accident Records of Four Types of Public Conveyances in the San Francisco Transit System for One Quarter

Type of conveyance	No. of accidents	Car hours operated	Car miles operated	Accidents per 10,000 hr. of operation		Accidents per 100,000 miles of operation	
				Actual no.	Relative no.	Actual no.	Relative no.
Motor coach	389	219,426	2,265,331	17.7	100	17.2	100
Trolley coach	54	21,933	204,550	24.6	139	26.4	153
Trolley cars	2,189	558,660	4,923,643	39.2	221	44.5	259
Cable cars	157	20,122	117,328	78.0	441	133.8	778

accident rates computed, on the basis of hours of operation and on the basis of miles of operation, for four types of public conveyances in the city of San Francisco. Although the relative accident rates of trolley coaches and trolley cars, as compared with that of motor coaches, do not differ greatly when accident rate is computed on the basis of miles or hours of operation, the difference is very marked for cable cars. When hours of operation are used as a base the accident rate of cable cars is seen to be something over four times as great as that of motor coaches,

but when miles of operation are used the accident rate of the cable cars is found to be nearly eight times as great. The particular base used in computing accident rate will necessarily vary with the type of job being considered and the nature of the problem.

RESULTS OF ACCIDENTS

Costs. Since accidents may involve damage to person, equipment, and property, they are expensive. In one transit system it was found that, taking into account all accidents whether major or minor, the average cost per accident in terms of claims paid was $50. In one year this organization paid out more than half a million dollars in claims. These figures do not take into account money required for repairs to the organization's equipment. In addition to these direct costs there are many hidden or indirect costs resulting from an accident. Heinrich has pointed out that, beyond claims paid and costs for repairs, there are further costs resulting from the lost time of the injured worker, the lost time by supervisors and other workers who investigate and assist in the accident, etc.[22] In one case cited by this author direct costs amounted to $209, and indirect costs amounted to $937. In another instance the accidents involved no direct costs, but indirect costs amounted to $154.

Effects on the Individual. It has been suggested that the individual who suffers an accident and recovers enjoys a certain amount of prestige among his fellows. However, if the accident is a relatively minor one, and the individual persists in describing it, he undoubtedly will lose as much prestige as do those persons who insist on telling about their operation long after the scar has healed. There is considerable clinical evidence indicating that if the accident is fairly severe the individual may develop rather marked personality disorders. He is likely to become apprehensive and irritable; neurotic symptoms may be induced. In extreme cases delusions of persecution may develop to the extent that the individual manifests vengeful reactions toward his employer or toward friends or members of his family.[44]

Social Consequences. The social consequences of accidents may be of considerable significance. The happiness, hopes, and sometimes even the health of the other members of the family are affected if an economic collapse results when the wage earner meets with a fatal or severe accident. Besides the effects on the family, many unpleasant social experiences may result from the recurring spells of despondency which are often characteristic of the permanently disabled and occupationally unfit individual.

It is generally believed that the ordinary workman is well protected

financially in the event of an accident. As Swan has pointed out, however, the extent of the protection is far from adequate.[39] At the present time in this country, all states but one require some kind of protection against occupational injuries or disease. Nevertheless the amounts of money received as a result of disability are generally insufficient to provide for the emergency. Furthermore, in most instances the individuals exposed to severe industrial hazards do not receive sufficient income to establish a satisfactory emergency reserve. Monetary benefits to the injured worker, of course, can never compensate for his suffering nor protect him from the destructive effects of continued idleness.

INDIVIDUAL DIFFERENCES IN SUSCEPTIBILITY TO ACCIDENTS

Distribution of Accidents. As with all types of measures of human performance, people differ with respect to the number of accidents that they incur in connection with their work. For the same period of time, some persons will not be involved in a single accident, *i.e.,* they will be accident-free, whereas others will be involved in one or more accidents. This fact of individual differences in accident rate has long been recognized. A distribution of accidents of the sort that is almost universally cited as being "typical" is shown in the top of Fig. 12-1. It will be noted from this table that the largest proportion of workers are accident-free, and that smaller and smaller numbers sustain one, two, three, etc., accidents respectively. Since distributions of this kind, which are described as *L-shaped curves,* are the ones that are customarily reported, the belief has grown up that they are characteristic of all types of accidents.

The shape of the distribution of accidents among workers would be of little more than academic interest were it not for the fact that an important and not wholly valid conclusion is drawn from them. From the L-shaped curve it follows that most of the accidents are attributable to a very few erring individuals. From this situation some investigators conclude that elimination of this small number of persons would markedly reduce the total number of accidents. If the foregoing is assumed to be true, then the major emphasis in a safety program should be given to locating and separating the so-called "accident repeaters." For example, in one industrial plant employing several thousand workers it was found that 20 per cent of the workers accounted for all accidents, but 6 per cent accounted for 65 per cent of all accidents.[36] Presumably, therefore, the elimination of this 6 per cent of the workers would reduce accidents by 65 per cent.

In generalizing from results such as these to all other situations where accidents occur, the following two assumptions must be made:

1. Provided there are no changes in methods and conditions of work, training procedures, etc., the accident rate for each individual must remain the same from one period of time to another, *i.e.*, each worker's accident rate is constant.

2. The distribution curves of all types of accidents, for all jobs, are L-shaped.

If the first proposition were not true, then in a period of time after the elimination of the few high-accident workers, others would move

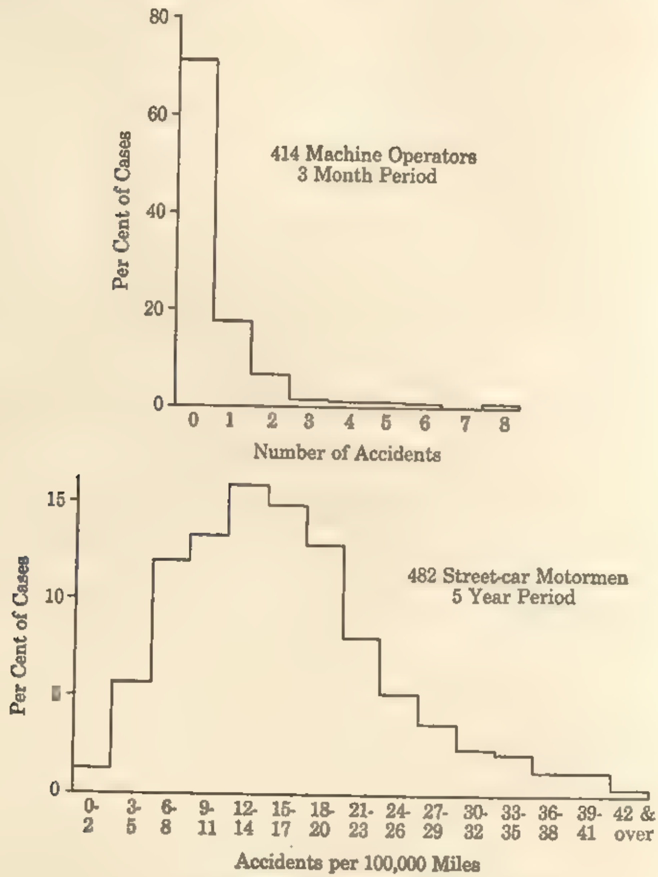
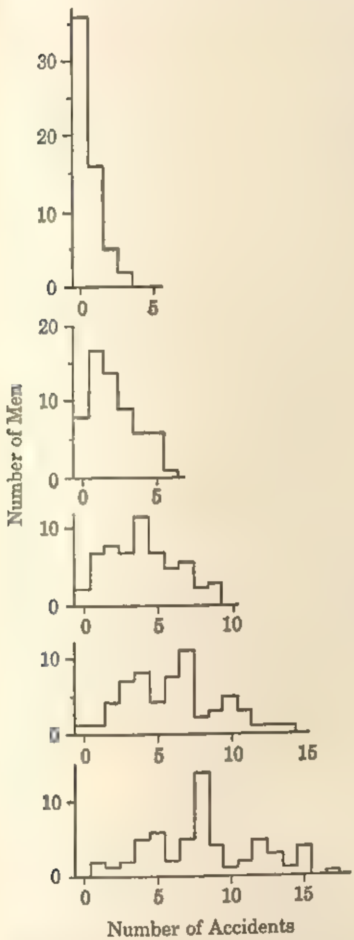


FIG. 12-1. Examples of distribution of accidents.

up to take their place. Furthermore, were these high-accident workers retained, their accident record might be considerably improved just in the natural course of events. Evidence pertinent to the inconstancy of accident rate will be examined in the next section. Suffice it to state here that accident rates are far from being constant; an individual's accident rate varies considerably from time to time, and the extent of this variation differs from one job to another.

In so far as the second assumption is concerned, there are many instances, particularly in the transportation industry, where the distribution of accidents is definitely not L-shaped. An example of a distribution of



Length of period during which accidents were recorded	Total number of accidents incurred	Per cent of accidents caused by poorest 10 per cent of operators
1 month	32	50
4 months	92	39
7 months	242	18
10 months	366	20
13 months	485	19

FIG. 12-2. Distribution of accidents of streetcar motormen, taken for various periods of time.

this sort is shown in the bottom of Fig. 12-1,[26] which can be compared with the distribution given on the top of the figure. L-shaped curves may be expected to occur under two conditions: first, when the job is a relatively safe one and the average accident rate is low, and secondly, when accident records are collected for only a very short period of time. The last condition is, of course, an artificial one. As longer and longer periods of time are used, the total number of accidents sustained by the group

of workers increases because of increased opportunity to have accidents, and the distribution of accidents departs more and more from the L-shaped curve. This is illustrated in Fig. 12-2, where the distributions of accidents for the same 59 trolley-car motormen are given for time periods of increasing length. It will be noted in this high-risk job that, for short periods of time such as 1 month, a small proportion of the workers account for a large proportion of the accidents and the distribution of accidents fits the conventional L-shaped curve. With longer periods of time, however, neither of these conditions is true.

Consistency in Accidents. As pointed out in earlier chapters, the measure of consistency of performance is given by the reliability coefficient. With accidents consistency in this type of behavior can be measured by observing the degree of correlation between accident rates for a group of individuals for two periods of time. In Table 12-3 are given the reliability

Table 12-3. Reliability Coefficients of Accidents for Several Groups of Workers

Reliability	No. of cases	Job	Time
.38, .44	524	Mechanic apprentices	Successive year periods [10]
.39, .36	259	Mechanic apprentices	Successive year periods [10]
.05, .27, .37	57	Dockyard apprentices	Successive year periods [10]
.33, .25	100	Dockyard apprentices	Successive year periods [10]
.57, .39, .37	94	Dockyard apprentices	Successive year periods [10]
.72	21	Lathe operators	Two 3-month periods [20]
.69	36	Lathe operators	Two 3-month periods [20]
.37	22	Profiling operators	Two 3-month periods [20]
.53	29	Profiling operators	Two 3-month periods [20]
.29	34	Bus drivers	8 months, odd and even months
.44	59	Trolley-car motormen	8 months, odd and even months
.30, .33, .21, .22	166	Bus drivers	Successive year periods [11]
−.08, .24, .02	86	Bus drivers	Successive year periods [11]
.07, .23, .30	67	Bus drivers	Successive year periods [11]
.18	398	Bus drivers	Successive year periods [11]

coefficients of accidents for several groups of workers. Experience indicates that, although the reliability of accidents frequently approaches zero, it hardly ever rises above .80. Compared to other measures of workers' performance, accident rate generally is much lower in consistency. This means that an individual is not highly consistent in his level of safety of performance over a period of time and therefore the

prediction of possible future accidents from past accident history will not be highly accurate.

Under present circumstances, it is impossible to account satisfactorily for the relatively low reliability of accidents. It may be that the tendency to have accidents is simply an unstable trait. On the other hand, the procedures for recording and describing industrial accidents are neither sufficiently accurate nor comprehensive. Perhaps more careful definition of what constitutes an accident and better descriptions of the events will lead to measures yielding higher reliability.

Accident Proneness. The term *accident proneness* has been used with so many different meanings that it can almost be considered useless. Sometimes it is used simply in a descriptive fashion to indicate that a particular individual has had many accidents. On other occasions the concept is used as an explanation of why people have accidents, referring to a syndrome of personality deficiencies. As a consequence, when the term accident proneness is used the ideas conveyed to a reader or a listener may be quite different from those intended.

In its most general meaning the term accident proneness refers to the tendency of persons to retain their relative liability to be involved in accidents. But even with this definition there is still confusion. The reference may be to the consistency of individual differences in accident rate of a given kind in a specific working situation. Thus the concern may be with the incidence of collision accidents of truck drivers. On the other hand, the definition is sufficiently loose that it may refer to the retention of liability to accidents of different types and to accidents under different circumstances. Here the notion is that accident proneness is a general trait that manifests itself in different types of accidents and in all situations. Thus a person who is unsafe as a truck driver would be expected to have many accidents if he became a bus driver, a lathe operator, or a food packer.

If it is held that accident proneness merely means consistency in performance with respect to specific circumstances, then the important evidence is the correlation between accident rates in two successive intervals. The degree of consistency with which individuals maintain their rate of accidents of a given kind over time was discussed in the previous section. It was shown that although individual consistency (reliability) is certainly not high, there is some tendency for an individual to maintain a particular liability.

The evidence with respect to the validity of the conception of accident proneness as a general trait of the individual can be found in studies of the degree of relationship existing between different types of accidents. Newbold studied the relationship between two types of accidents in-

curred by groups of factory workers, and their relationships to accidents that occurred at home.[31] The coefficients of correlation obtained were of the order of .20 to .30. Farmer and Chambers investigated the relationships among different types of accidents of bus drivers, the accidents being classified in terms of cause such as errors of judgment, skids, etc.[11] On the average the coefficients of correlation were of the order of .10, and ranged from −.02 to .39. Correlations of the same order between collision and noncollision accidents have been reported for bus drivers and streetcar motormen by Brown and Ghiselli.[3,4] All of these correlations are fairly low and indicate little generality for liability to have accidents of different kinds and in different situations. Hence it may be concluded that if there is a general trait of accident proneness it is a very unimportant factor in the determination of accidents.[28]

THE PREDICTION OF INDIVIDUAL LIABILITY TO ACCIDENTS BY PSYCHOLOGICAL TESTS

While investigators concerned with industrial safety problems have long emphasized the importance of the selection of workers in terms of liability to have accidents, there have been surprisingly few systematic investigations of the usefulness of psychological tests in the selection of safe workers. For one thing, in most industrial organizations, problems of personnel and problems of safety are handled by departments having little or no connection. The personnel department, which is more closely tied to operations and production, has tended to evaluate selection devices in terms of the degree to which they predict output rather than safe operation. Thus in most instances the safety department, because of organizational structure, is prevented from stating its case when new employees are being selected.

A second factor of importance is the inadequacy of accident records. In general most industrial accident records either are incomplete or are in such a form that it is impossible to ascribe the accident to the worker concerned or to other persons or agencies. Under the circumstances it is not surprising to find that most of the studies on the use of tests in the prediction of accidents have been made in the transportation industry, where accidents are of prime importance both from the point of view of costs and of smooth operation.

Tests in Relation to Safety in Industrial Operations. Results obtained by Farmer and Chambers from the application of a variety of tests to three groups of young apprentices showed that no single test gives satisfactory results.[12] A weighted combination of the more valid tests, however, gave a satisfactory correlation with accidents, a coefficient of the order of .45. Certain of these tests were administered to other groups

of workers. The validity coefficients of these tests are given in Table 12-4. Again, although the individual tests gave unsatisfactory results, weighted combinations of them predicted accidents to a reasonably acceptable degree.

Table 12-4. Validity of Various Tests for the Prediction of Accidents of Workers in Several Different Occupations

Test	Shipwrights	Electrical fitters	Engine fitters	Naval artificers I	Naval artificers II
Reaction time	0.20	−0.07	−0.16		
Dotting	0.25	0.13	−0.06	−0.08	0.14
Pursuit	0.18	0.16	0.16		
Number setting	0.27	−0.15	−0.10		
Intelligence	0.12	−0.01	0.31	0.09	−0.03

Schaefer studied the relationship between intelligence test scores and accident rates of 6,829 industrial workers.[33] His results are shown in Table 12-5. For these workers there is definite positive relationship between test scores and accidents. The variation in accident rate for workers at any given score level is not indicated, however, and therefore the magnitude of the relationship between test scores and accident rate cannot be estimated. Henig also has reported a positive relationship between intelligence test scores and accidents for trade school students.[23]

Table 12-5. The Relationship between Intelligence Test Scores of Industrial Workers and Accident Rate

Intelligence test score	*Accidents per year per 100 workers*
(High) A	0.66
B	1.12
C	1.53
D	2.08
(Low) E	2.76

Tests in Relation to Safety in Transportation Operations. In the transit industry Viteles made the first major study in this country of the predictability of accidents by test scores.[47] He devised a multiple stimulus-reaction test, which required the individual to make different types of

coordinated responses to different kinds of stimuli. For a group of motormen, a coefficient of correlation of the order of .40 was obtained between scores on this test and supervisors' ratings of safety. On the other hand, two intelligence tests were found to be unrelated to this criterion of safe operation. A further check on the validity of Viteles' multiple stimulus-reaction test was made by Shellow in which test scores were compared with actual accident records.[34] Although still showing the test to be useful, the findings indicated a somewhat lower validity. Slocombe and Brakeman employed a test for streetcar motormen that involved accuracy and speed of movement.[38] Again, although this test showed some validity, it was of rather low degree. They studied 92 operators, 54 per cent of whom were considered low-accident men. The proportions of low-, medium-, and high-scoring men found in this group of low-accident cases were as follows:

	Per cent
Low scorers	56
Medium scorers	53
High scorers	50

Brown and Ghiselli, in studying a group of bus drivers, found a coefficient of correlation of only .05 between intelligence test scores and accidents.[3] Other investigators have similarly found that intelligence tests are of no value in predicting safety of operation in the transportation industry.[11]

Table 12-6. Validity Coefficients of Various Tests in the Prediction of Accidents for an 8-month Period for Three Groups of Streetcar Motormen

Years of experience	1 to 2	3 to 10	15 to 30
Number of cases	33	53	47
Type of test			
Tapping	.41	.01	.13
Dotting	.32	−.05	.10
Foot-reaction time	.36	.21	.07
Steadiness	.25	.16	−.03
Visual acuity	−.09	−.04	.29
Field of vision	.08	.09	.34
Glare vision	.07	.19	.10
Distance judgment	.00	.17	.19
Judgment of distance	.18	−.01	.23
Distance discrimination	.24	.10	.10
Mechanical principles	.15	−.01	.22

Minium administered a variety of tests to three groups of streetcar motormen varying in length of service on the job and compared the scores with accident records for an 8-month period.[27] The results, shown in Table 12-6, show little consistency from one experience group to another. Inconsistency of findings are fairly common and in part may be ascribed to unreliability of accidents. But in addition it suggests that prediction of safety of performance by means of tests is quite specific to the job and to the types of workers involved. Minium's results indicate that tests of speed and accuracy and of judgment of distance give the best predictions. These findings were verified for a group of taxicab drivers by Ghiselli and Brown.[18] Their findings, presented in Table 12-7, also point to the value of interest tests.

Table 12-7. The Validity of Various Tests in the Prediction of Accidents of Taxicab Drivers

Type of test	*Validity coefficient*
Dotting	.35
Tapping	.47
Judgment of distance	.18
Distance discrimination	.20
Mechanical principles	.11
Arithmetic	−.09
Speed of reactions	−.04
Interest	.28

The National Safety Council has reported results on a series of tests administered to 180 motor-coach operators.[30] On each test a "passing" score was set as the average score made by operators who had fewer than the average number of accidents. The results, given in Table 12-8, suggest that certain types of simple motor tests may have some validity. The stability test, involving bodily sway in the standing position, and the muscular-reaction test, involving measures of balance, coordination, relaxation, and tremor, show a positive but low correlation with accident rate. Reaction time also is related to accident rate, the quick reactors having the higher rates. The visual tests show practically no relationship with safety of operation. In connection with reaction time tests it is interesting to note that Wechsler, in studying taxicab drivers and employing a test involving various responses to a series of stimuli, found that both very quick and very slow responses were associated with poor accident records.[49] The number of errors made on Wechsler's test was positively correlated with accidents.

Table 12-8. The Relationships between Scores on Several Simple Tests and Accidents of Motor-coach Operators

Accidents in 9-month period	Operators "passing" test, %					
	Bodily stability	Simple reaction time	Muscular reaction	Depth perception	Field of vision	
					Right eye	Left eye
0–1	71	57	100	86	72	86
2–3	90	70	90	80	100	100
4–7	65	72	83	87	85	84
8–9	48	86	76	76	90	86
10–12	55	86	91	82	59	77
13–17	45	89	66	89	78	89

Evaluation of Selection Tests with Respect to Accidents. The foregoing review relative to the usefulness of psychological tests in the prediction of accidents, although not exhaustive, nevertheless presents typical findings. It is apparent that weighted combinations of tests may give reasonably good results, but there are no particular types of tests that are uniformly outstanding. The generally low validity of tests in the prediction of accidents has been somewhat disappointing. Before it is finally concluded that prediction by tests is inadequate, however, the nature of the characteristic being predicted should be more closely examined. It may well be that the fault does not lie in the tests but rather in the measure of performance, viz., reported accidents, with which the tests are being correlated.

It was noted in a previous section that the reliability of accidents, as they are recorded in the industrial situation, ordinarily is not high. This means that accidents are not particularly good predictors of themselves. Under these circumstances it can hardly be expected that behavior of another kind, *i.e.*, performance on a test, can give a wholly satisfactory prediction. When considered in the light of the unreliability of accidents, the predictive effectiveness of tests appears considerably better than judged at first glance. If the index of accident liability could be made more reliable, the predictive effectiveness of tests would probably be increased. If only those accidents that were due solely to faulty per-

formance on the part of the worker were considered, then perhaps prediction would be more effective.

THE INFLUENCE OF WORKING CONDITIONS

Temperature. In a study of the relationship between temperature of the work environment and frequency of minor accidents, Osborne and Vernon found for groups of factory workers that accident rate was at a minimum during the more moderate temperatures.[32] Their results, graphed in Fig. 12-3, show that as there is either an increase or a decrease from

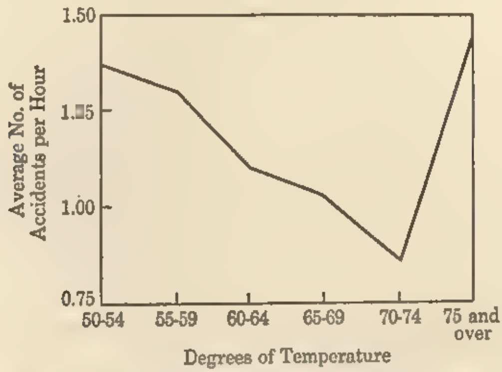


FIG. 12-3. The relationship between temperature in the workshop and accidents of factory workers.

moderate temperatures, accident rate is increased. These findings are typical in that they point to an optimal temperature at which accidents are at a minimum, and that with either warmer or colder atmospheres accidents occur with greater frequency. There is some indication that with extreme temperatures, not only is accident rate increased, but also the severity of accidents is greater.[46] Thus, when a worker suffers an accident under conditions of higher temperatures, the accident is likely to be of a severity that will cause him to be off the job for a longer period than with an accident incurred under more moderate temperature conditions.

It is of interest to note that the effects of extreme temperatures are more marked with older than with younger workers.[46] Under conditions of very high temperatures the accident rate of the older workers is much higher than that of the younger ones, whereas with more moderate temperatures the differences between the older and younger workers are less significant.

Illumination. As might be expected, accident rate shows an increase under conditions of poor illumination. In one survey of a variety of different occupations it was found that, under conditions of inadequate artificial lighting, accident rate was 25 per cent greater than under ordinary daylight illumination.[44] Under inadequate illumination certain types of accidents show a greater increase than other types. In the survey just mentioned it was found that under poorer conditions there was a greater increase in accidents due to workers falling than in accidents due to either machinery or falling bodies. It has been noted in factories that, at the onset of darkness toward the end of the working day before the plants are completely lighted by means of artificial illumination, the incidence of accidents shows an increase.[19]

Day Shift versus Night Shift. Allied to the problem of illumination is that of the effects of night work upon accidents. Vernon has noted that in four different munitions plants the accident rate among workers on the night shift was 17 per cent less than that among workers on the day shift.[43] The evidence indicates that, in part at least, this superiority of the night shift is due to the more effective illumination provided by artificial means as compared with that afforded by natural means during the day shift.

Vernon studied the frequency of accidents on the night shift, in relation to adequacy of artificial illumination.[43] Accidents of the night shift were expressed as a per cent of those of the day shift. The findings are presented in Table 12-9. It is obvious that on the night shift accidents are

Table 12-9. Accident Rate of Factory Workers on the Night Shift in Relation to Adequacy of Illumination—Accidents Expressed in Terms of the Accident Rate of Day-shift Workers

(Day-shift accident rate = 100)

Type of factory operation	Condition of lighting	All accidents	Eye accidents
6-in. shells	Good	80	94
9.2-in. shells	Good	81	99
9.2- and 15-in. shells	Moderately good	81	103
Fuses	Fairly good	90	141

relatively more frequent under poorer conditions of artificial illumination than under conditions of good illumination. The table also shows that with eye accidents, as compared with other types of accidents, the night

shift stands in a relatively poorer position. It is apparent that under conditions of adequate illumination, accident rate on the night shift is likely to be less than that on the day shift. Furthermore, the relative superiority of the night shift is not the same for all types of accidents.

FATIGUE AND WORK

Fatigue. As has been pointed out earlier, the concept of fatigue is an ambiguous one. The term fatigue does not describe a single state of the individual but covers several different states of the individual that may vary independently. Thus at a given time the individual might be considered to be fatigued in one sense but not fatigued in another. Any explanation of the performance of industrial workers in terms of fatigue will therefore be unsatisfactory.

Circular reasoning with regard to the concept of fatigue is particularly evident when fatigue is offered as a cause of accidents. If a single instance is found in which fatigue is present, it is used as positive proof that fatigue causes accidents, but those instances in which contrary results are found are not considered as negative evidence, *i.e.*, as denying that fatigue causes accidents. In the latter cases it is argued that fatigue was just not operating. Suppose that in one factory it is found that accidents increase toward the end of the working day. It is then argued that this is evidence that the fatigued worker is more likely to have accidents. On the other hand, if there is no increase in accidents during the latter part of the day, it is said that the work was simply not fatiguing; and if accidents actually decrease throughout the day the explanation is offered that totally different factors, such as reduced opportunity to have accidents, are at work.

It is to be noted that these three sets of findings can just as readily be used as evidence that fatigue does not produce an increase in accidents. The last two facts could be used as evidence that the fatigued worker does not have a heightened accident rate, and the first fact, that accidents increase toward the end of the day, could be explained away by saying that it is due to factors other than fatigue. It does not appear advantageous, therefore, to consider accidents in relation to such a nebulous concept as fatigue. It will be more fruitful to consider the nature and conditions of work as they are related to accident rate. In the following discussion, therefore, consideration will be given to diurnal variations, speed of work, severity of work, and length of the work period in relation to accidents.

Diurnal Variations in Accidents. In Fig. 12-4 is shown the frequency of accidents in a factory in relation to daily hours of work for both the day shift and the night shift.[10] These curves appear to be fairly typical

for industrial work.[44] The frequency of accidents in both the morning and afternoon periods gradually increases and then decreases toward the end of the working spell. With day-shift workers the frequency of accidents in the morning does not appear to be significantly different from that in the afternoon. For the night shift, accidents generally are considerably higher in the first working spell than in the second. It is to be understood of course that the shape of these diurnal curves varies with the nature and conditions of work.

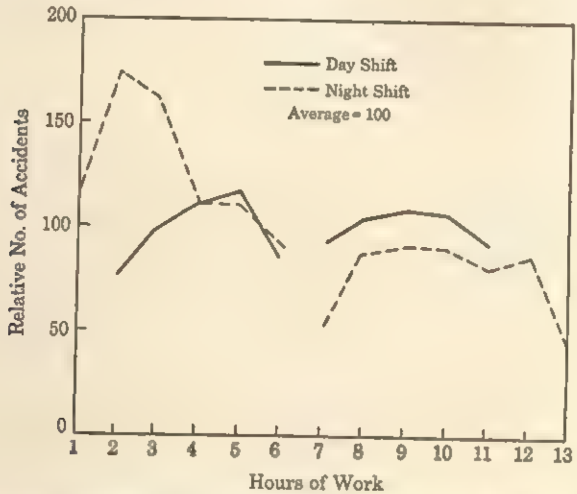


FIG. 12-4. Diurnal variations in frequency of occurrence of accidents of factory workers on the day shift and on the night shift.

Speed of Work in Relation to Accidents. If the daily curves of output are recalled, it will be noted that they are similar in shape to those of accidents. This has given rise to the suggestion that industrial accidents are a function of speed of production; *i.e.*, during those hours when the worker operates at a slower rate the frequency of accidents will be less than during the hours when he operates at higher speeds.[42] If this were true, then the number of accidents per unit of output should be constant throughout the day.

The notion that speed of work is a major determinant of accidents has led to considerable controversy. Osborne and Vernon found that for different groups of factory workers on the day shift the ratio of accidents to output was relatively constant with the exception of the final hours of work, when it was somewhat reduced.[32] For similar workers on the night shift, however, the ratio of accidents to production was found to decrease gradually. Illustrative findings of these investigators are given in Fig. 12-5. On the other hand, results obtained by Goldmark, Hopkins, and

Florence, shown in Fig. 12-6, indicate a rise in the number of accidents per unit of production from the beginning to the end of the working day.[19] Since the results of studies concerning the relation of the speed of work at different intervals of the working day and accident rate are not in agreement, no final conclusion is justified at this time.

At first thought it might be expected that the frequency of accidents would vary directly with the speed of work. This expectation has two different bases. First, when the individual is working faster he has in-

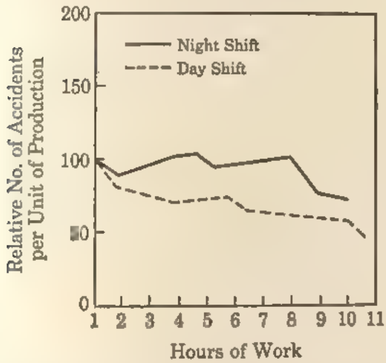


FIG. 12-5. Relative number of accidents per unit of production of day-shift and night-shift factory workers in relation to daily hours of work.

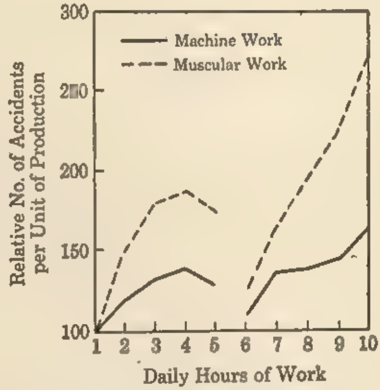


FIG. 12-6. Relative number of accidents per unit of production of physically demanding (muscular) factory work and less demanding (machine) work in relation to daily hours of work.

creased opportunity to have accidents, and secondly, under conditions of increased speed of work he is not able to exercise the care that he can when working more slowly. Critical examination of these two points will throw some doubt upon the notion that in industrial situations accident rate is directly conditioned by speed of work.

The first proposition hinges directly on the nature of the work. If the hazards that lead to accidents arise directly out of the industrial operation itself, and if they have no connection with surrounding conditions or other aspects of the process, then it would be true that each unit of work would provide the same opportunity to have accidents. Only when an individual is operating his equipment could he have an accident. When an individual is working faster it simply means that the machine is operating more continuously, and the periods between operations when there is no exposure to accidents are fewer in number. In a working situation such as this, the argument that accidents are a function of speed of work would have to be accepted. With other kinds of jobs the hazards

might not arise from the individual's work but rather from surrounding activities. In these jobs, speed of production would not be highly related with accident rate. Obviously most industrial work hazards will arise from both the work itself and from surrounding activities. The degree to which accidents will be a function of speed of work will therefore be determined by the nature of the particular situation.

The second argument is rather tenuous, owing to the lack of pertinent data. It is clear that with excessive speeds of operation the individual will not have opportunity to foresee and avoid events leading to accidents. For example, persons who drive cars at high rates of speed have more accidents than those who drive at slower speeds.[8] Experiments in the psychological laboratory on quality of performance show speed of work to be an important factor.[20] Increasing the speed of work produces a disproportionately greater increase in errors. In other words, as speed of work is increased from very slow to very high speeds, there will be no great change in errors at first and then they will begin to increase markedly.

The pertinent question to ask is: Just what range of speeds is under consideration? Certainly it would seem likely that in situations where the worker must continue on his job for long periods, operations cannot proceed at excessive speeds, *i.e.*, at a rate to be found in the area of the curve where errors begin to mount significantly. Since there are no data describing the speeds of industrial tasks, and since the critical speeds doubtlessly will vary from one task to another, no satisfactory answer can be given to this question.

The foregoing arguments are not to be interpreted to mean that the ratio of the number of accidents to the amount of production is a meaningless figure. Rather, as was indicated in the earlier discussion of accidents, amount of work constitutes a useful base for comparing accidents occurring under different circumstances.

Severity of Work in Relation to Accidents. By severity of work is meant the extent to which the operations make physical demands upon the worker. Some jobs, such as those in the clerical occupations, require far less physical effort in their performance than others, such as jobs in the industrial occupations. In terms of the mental effort demanded by the work, of course, the situation may be reversed. Although there are no precise indices for indicating the physical demands of jobs, it is possible to group jobs roughly into classes according to physical demands in order to make crude comparisons.

The evidence concerning accidents in relation to severity of work is not as complete as might be desired. Direct comparison of accidents incurred in more demanding and less demanding types of work is difficult since the nature of the work will be different. It does appear, at least by

some indices, that physically demanding work is less safe than less demanding work. No great differences were found when comparing the frequency of accidents in the afternoon spell with that in the morning for more and for less physically demanding jobs. Goldmark, Hopkins, and Florence found that for factory workers engaged in muscular work the ratio of afternoon to morning accidents was not significantly different from that for persons engaged in machine or handwork.[19] The afternoon productivity of the men engaged in the heavier work, however, was relatively less than that for the men engaged in the lighter work. In terms of accidents per unit of production, therefore, the workers on the more demanding jobs appeared to be at a disadvantage. The diurnal curves for two of these groups, in terms of accidents per unit of work, are given in Fig. 12-6. It will be noted that, as the working day proceeds, the number of accidents per unit of production shows a significantly greater increase for men in jobs requiring more physical effort.

Length of the Work Period. It is generally presumed that increasing the length of the working period will result in a disproportionate increase in accident rate. The results do not bear out such a notion. To be sure, with longer periods of work the individual has greater opportunity to have accidents in terms of time, but if his rate of work is slower, the number of accidents per unit of work may not show any increase.

Osborne and Vernon found that for groups of factory workers a reduction of the length of the work period did reduce accidents.[32] For a group of workers on the day shift, reduction of the working week from 59 to 35¼ hr.—a reduction in working time of 40 per cent—brought about a reduction in accidents of 10 per cent. For night-shift workers, reduction of the weekly hours from 63 to 48 hr.—a reduction of 24 per cent—brought about a reduction in accidents of 25 per cent. It is apparent from these findings that with industrial workers the frequency of accidents is related to the length of the work period, but the particular nature of the relationship is not clear.

Somewhat more illuminating results obtained by Vernon are given in Table 12-10.[42] Here frequency of accidents is related to the length of work period and to the speed of work. For the industrial workers represented in this investigation it can be seen that lengthening the work period produces an increase in frequency of accidents roughly proportional to the increase in time. Thus longer periods of time might be considered less safe than shorter periods. However, with longer periods of work, speed of work is decreased. Thus, when number of accidents is related to units of work turned out, it appears that longer periods of work appear to be more safe.

The nature of the relationship between hours of work and safety quite

Table 12-10. Accidents of Factory Workers in Relation to Length of the Work Period and Speed of Work

	Hours of work per week	Relative hours of work per week	Relative no. of accidents per week	Relative speed of production per hour	Relative no. of accidents per unit of production
Men	44.9	100	100	100	100
	46.8	104	106	98	82
	51.8	115	114	92	69
Women	48.0	100	100	100	100
	51.4	107	111	88	99

naturally will vary with the nature and conditions of work. Obviously, with production-line operatives and others whose speed of work is set by machines and schedules, the association expected between hours of work and accident rate will be different than the relations described above. But even in such cases it has yet to be shown that longer hours produce a disproportionate increase in accidents.

There is relatively little sound information concerning hours of work and safety of performance in the transportation industry. Such information as is available is largely negative, indicating that hours of work are unrelated to accidents. Minium compared the hours worked per year by streetcar motormen with the number of accidents in which they were involved.[27] The number of hours these men were expected to work was normally 2,400, but some of them worked as little as 1,200 hours, while others through overtime put in as many as 3,600 hours. The findings indicated that the number of accidents incurred by the men during the year was unrelated to the number of hours worked. A study made during World War II of the landing accidents of bomber aircraft revealed no consistent relationship between duration of the mission flown and safety.[25]

PERSONAL FACTORS IN ACCIDENTS

Experience in Relation to Accidents. Many reports have been published that presumably show a relationship between experience and accidents.[44] In general these findings indicate that with increasing amounts of experience, accident rate shows an initial rapid decrease followed by a more gradual reduction. Typical results of investigations of this kind are given for streetcar motormen in Fig. 12-7. All these studies suffer from

an important error in so far as the assessment of the effects of experience is concerned. The procedure in these investigations is to divide the workers in a given organization into experience groups and then calculate the accident rate of each group. This method must necessarily assume that, if there were no differences in experience, all these different groups of workers would have the same average number of accidents. It is obvious, however, that in many jobs high-accident men will be retired through injury, will be separated, or will leave the employment voluntarily. This

FIG. 12-7. Accident rate of streetcar motormen in relation to years of experience on the job.

means that those persons who stay on the job for longer periods of time will be selected in terms of safety in operations. The drop in accidents with increased time on the job then may be due, in part or in whole, simply to the dropping out of the high-accident operators. With such data the effects of experience and the effects of this biased selection are so intimately related that they cannot be distinguished.

In order to evaluate the effects of experience it is necessary to follow the accident history of the same group of workers over a long period of time. In this manner the effects of selective elimination of workers will be avoided. In Fig. 12-8 are shown the accident rates per month of a group of streetcar motormen for the first 19 months of their employment.[17] By the end of the fifth month on the job their accident rate was halved, and by the end of the nineteenth month it was reduced to about one-third. A relationship between accidents and experience is thus demonstrated.

To evaluate the effects of experience over a period of years is quite difficult. Conditions and methods of work, together with type of equip-

ment, are likely to change, thus rendering accidents in one period not comparable with those in another period.

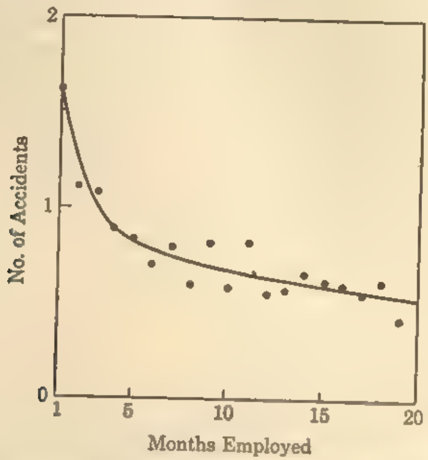


FIG. 12-8. Relationship between months of experience of streetcar motormen and accident rate.

Experience is not equally effective in reducing all types of accidents. Some types of accidents are much more easily avoided or eliminated through learning on the job than are other types. In Table 12-11 are

Table 12-11. Relative Numbers of Accidents of Various Types Incurred by Bus Drivers in Relation to Years of Experience

Cause of accident	Years of experience				
	1	2	3	4	5
Errors of judgment	100	57	57	50	41
Over-runs	100	78	88	106	91
Skids	100	45	14	19	7
Miscellaneous	100	94	79	89	57

given illustrative findings for bus drivers presented by Farmer and Chambers.[11] Here it will be seen that accidents resulting from skids are much more amenable to reduction through experience than are those resulting from over-runs.

Age in Relation to Accidents. The situation with regard to age as a factor in accidents is similar to that of experience. In typical investiga-

tions of the relationship between age and accidents workers are simply divided into age groups, and the accident rate for each group is computed. In most instances the results show that accident rate decreases with increasing age, although this relationship is by no means universal.

In almost all studies age is highly correlated with experience, and therefore any relationship obtained with accidents cannot definitely be ascribed to either. Certain attempts have been made to hold experience constant by statistical devices. In one such investigation Newbold showed that when experience of factory workers was held constant by statistical means, the relationship between age and accidents was fairly high and negative.[31] Such relationships, however, do not hold for all jobs. With streetcar motormen Minium found that when experience was held constant by statistical means the relationship between accidents and age was practically zero, but this may not be wholly adequate since the more experienced workers are likely to be those who remained on the job because of their safe performance.[27] It is probable that the selective factors operating in these cases prevent strict compliance with the basic assumptions inherent in the statistical method.

One fact that does seem clear is that when an older worker suffers an accident the consequences are likely to be more severe than they would be for a younger worker. Vernon, Bedford, and Warner found that when older coal miners suffered accidents they were likely to be off the job for longer periods of time.[45,46] An example of their findings is given in Table 12-12. Similarly in the steel and iron industry, Chaney and Hanna

Table 12-12. The Relationship between Age and Severity of Accidents among Coal Miners

Days absent from job following accident	Under 30 years	30 to 39 years	40 to 49 years	50 years or more
	Per cent			
Under 20	67	65	56	48
20 to 59	28	28	35	41
60 and more	5	7	9	11

found that the accidents suffered by older workers were more likely to result in death or permanent disability.[5]

Health and Physical Defects in Relation to Accidents. As might be expected, not only is general health related to proneness to have acci-

dents, but also physical defects of various sorts predispose the individual to increased accident liability. In a study of nearly fifteen thousand trades and dockyard apprentices, Farmer and Chambers found a correlation of the order of .30 between accident rate and number of illnesses.[10] A relationship of about the same degree was obtained by Newbold for workers in various manufacturing industries.[31] Viteles compared the accident rate of electrical substation operators who had physical defects with that of operators who did not have such defects and found that the number of accidents of the former was three times as high as that of the latter.[48] Slocombe and Bingham, in a study of streetcar operators over fifty years of age, found that the accident rate of men with abnormal blood pressure was more than twice as high as that of men with normal blood pressure.[37] Among industrial workers, Tiffen has compared the accident rate of workers who failed various visual tests with that of workers who passed such tests.[40] His results, given in Table 12-13, show

Table 12-13. Accidents of Industrial Workers in Relation to Their Performance on Visual Tests

Test	% having lost-time accidents	
	Passing test	Failing test
Acuity, far	1.09	1.37
Depth perception	1.23	1.57
Color vision	1.21	1.14
Phoria	1.26	1.82

that deficiencies in acuity, distance perception, and phoria (distance vision) are associated with accidents but the degree of relationship is small.

The foregoing evidence should not be interpreted to mean that the physically handicapped individual has no place in industry. In many instances the worker is unaware of his deficiencies and is therefore unable to compensate for them. When he becomes cognizant of his limitations and is instructed in compensatory procedures, the job situation can be made more favorable. Furthermore, employers do not always realize the importance of physical limitations and fail to place workers in suitable jobs. With an accurate program of placement, in which the physical demands of the job are accurately known and are compared with the physical limitations of the worker, accidents can be minimized. Hanman

compared the accident rate of a group of physically handicapped mechanical and manual workers in a Navy depot with that of a group of able-bodied workers.[21] He found that, with adequate placement of the physically limited individuals, their accident rate was even less than that of the able-bodied workers.

Personality and Emotional Factors in Relation to Accidents. There is abundant evidence that the personality and emotional states of the individual are important factors in accidents. Hersey, in a comparison of the mood of industrial workers with their accidents, found that 50 per cent of their accidents occurred during the period when they were emotionally depressed, such periods of depression comprising only about 20 per cent of their working time.[24] In a study of streetcar motormen it was found that in one-fifth of the accidents personality maladjustment or improper mental attitude was the primary factor.[48]

The importance of such findings is not to be minimized, since they do point to the role played by emotional factors; nevertheless it should be noted that the types of behavior referred to are symptoms of psychological disturbance rather than psychological determinants of accidents. In order to understand the emotional factors in accidents, it is necessary to direct attention to the emotional conditions contributing to unsafe performance rather than to be content with describing behavior that is associated with such performance. It is of little help, for example, to offer uncooperative behavior, inattention, and poor attitude as explanations of the causes of accidents. Only in a superficial sense are these factors productive of accidents.

What is necessary is some understanding of the factors that produce not only uncooperative behavior, inattention, and poor attitude, but also accidents. These unsatisfactory ways of behaving may have very different bases in different individuals. One individual may not have the so-called "safety attitude" because he has had trouble with his superior and, at the moment, does not like his job or anything connected with it. Another worker may be similarly described because he is unable to concentrate on his job due to financial or home worries. The situation cannot be corrected merely by emphasizing a change in attitude. The personality and emotional factors that underlie the poor attitude and at the same time condition the individual toward unsafe behavior must be uncovered and dealt with.

From the foregoing discussion it is not to be implied that personality and emotional factors are considered as if they were the direct implementing factors in accidents. Such a notion would not be true except, perhaps, in cases of sabotage or as manifestations of sadistic or suicidal tendencies. These factors should be considered as predisposing conditions likely to lead to accidents. The individual who is emotionally dis-

turbed will not necessarily have more accidents than a well-adjusted person; rather, he is *more likely* to have a higher accident rate.

The sources of emotional maladjustment predisposing the worker to accidents are many and varied. In some instances they are job-connected. Inadequate relations between a worker and his supervisor are likely to result in such disturbances. Fear of machines, worries about continuing on the job, etc., may produce emotional upsets leading to accidents. In other instances the emotional disturbance may arise from home, family, and financial conditions. Thus the individual involved in marital difficulties may be so depressed that his performance on the job is affected. Some emotional depressions are only temporary or transitory in nature. For example, the bus driver who is normally a safe operator might be unable to concentrate on the job when his child is seriously ill. Other unfavorable emotional states are of longer term. The individual with chronic financial worries might show a high accident rate over an extended period of time.

From clinical studies of individual cases some evidence is found which suggests that as a group those individuals who become involved in many accidents manifest less emotional stability than those who are accident-free.[1,9] Such people tend to conflict with authority, to be fatalistic in outlook, and to display anxiety. In addition, persons with bad safety records tend to have poor social and occupational histories.[50] However it is to be noted that the relationships between personality characteristics and accidents found in these clinical studies are rather low and are only observable when extreme groups are compared.

SAFETY PROCEDURES

Analysis of Conditions and Causes of Accidents. In any accident-prevention program the initial step is fact-finding. In order to be in the best position to develop preventive measures, a clear and detailed description of the present picture with regard to accidents is essential. Two kinds of information are desirable, viz., data relative to the type of accidents that occur and data relative to the conditions under which these accidents occur.

In the section concerned with the definition of accidents it was pointed out that in classifying accidents according to types three different systems could be used. Accidents can be classified according to the nature of the event, the causes of the event, and the effects of the event. It is important that these three ways of classifying accidents should be borne in mind. The descriptive categories set up for recording detailed facts about accidents should be homogeneous in relation to the particular classification system used. That is, the categories of the classification system should be

wholly concerned with the nature of the accidents, or the causes, or the effects; the different types of categories should not be mixed indiscriminately.

Certain examples of the three kinds of classification can be cited in order to illustrate their use. In Table 12-14 are given the accidents of

Table 12-14. Accidents of Streetcar Motormen Classified According to the Nature of the Accident

Nature of accident	Average no. of accidents per motorman per month		% of accidents	
Collisions	.745		95	
Collisions with automobiles		.680		87
Collisions with other streetcars		.025		3
Collisions with pedestrians		.040		5
Noncollision accidents	.038		5	
Boarding		.023		3
On board		.015		2
Total	.783		100	

streetcar motormen classified according to the nature of the event. It is apparent from this table that the major problems were collision accidents, particularly collisions with automobiles. These findings suggest that selective devices measuring such aptitudes as division of attention, perception of spatial relations, and judgments of speed would be useful in selecting potentially safe streetcar motormen.

In classifying accidents according to cause, the most general classification is a division of accidents into those attributable to human error and those attributable to material failure. For industrial accidents it is usually found that some 80 per cent to 90 per cent can be attributed to human error as compared to 10 per cent to 20 per cent attributable to inadequacies of machines and equipment. This means that preventive measures should emphasize the human factor. It is to be noted, however, that human error does not account for this large a proportion in all cases. With certain types of aircraft accidents it has been found that in instances where the cause could be definitely established, human error accounted for only two-thirds of the accidents. In other instances human error was found to be the cause in only 50 per cent of the accidents. Thus there is still a need for giving considerable attention to the material factors.

Investigations of accidents in terms of their effects gives information relevant to protective measures. Knowing the types of personal injuries and machine breakdowns and damages occurring, it is more likely that means can be found to prevent similar occurrences in the future. Among industrial workers, for example, injuries to fingers and eyes are by far the most frequent. Indeed injuries to the various parts of the arms and head account for almost three accidents out of every four. The prime problem obviously lies in the prevention of accidents of this sort.

In addition to information concerning the types of accidents that occur, it is desirable to know something about the conditions under which they occur. Knowledge of these conditions will be of aid in determining the preventive measures that can be most effectively applied. In previous sections various conditions, environmental and otherwise, which are associated with increased accidents have been touched on. In an attack on safety problems relative to any particular job, general knowledge of this kind will be helpful principally in providing an orientation for more detailed analysis. As already noted, there are rather marked differences between jobs, and therefore, in a particular job situation, data specific to that situation should be obtained.

As an illustration of the way in which information of this kind can be useful in accident prevention, a study of accidents in a printing shop can be cited.[14] Tabulation of accidents by type of press used was revealing of a particular source of trouble. It was found that the accident rate was about ten times higher on a particular type of press requiring somewhat unique movements than on other presses requiring more conventional movements. A further study of the situation indicated that for most machines the press came down slowly and picked off quickly. On the machine in question the speeds of movement were just the opposite. An analysis suggested that the high accident rate was due in part to the fact that the operators were not accustomed to a press with a rapid downward movement and a slow pickup. The movement of this particular machine was not consonant with the natural movements of the operators. It seemed therefore that this machine was inappropriately designed.

Design of Equipment in Relation to Safety. While there have been a number of experimental investigations conducted with respect to the design of equipment, most of them have been centered upon discovering ways for improving amount and accuracy of work. Little thinking has been directed toward the establishment of a set of principles concerning design of equipment for safety of operation yet a machine designed to yield the most effective production is not necessarily the one which can be operated with the greatest safety.

It is apparent, of course, that the mechanical design of a piece of equipment should be related to the physiological and psychological

characteristics of the operator.[41] This means that machines and working areas should be built around the operator rather than being designed solely for mechanical and functional effectiveness. But statements such as these are too general to be of help in dealing with specific problems. On purely rational grounds it would appear that there are two guiding principles to be followed in the design of equipment for safe operation, namely, the equipment should be designed to minimize the possibility of the individual getting into trouble, and to maximize the possibility of his freeing himself from difficulty once he is in trouble.

Several implications can be drawn from the principle of designing to minimize the possibility of the individual getting into trouble. First, the worker should be prevented from placing himself in a dangerous situation. Industrial engineers for several decades have been designing and installing guards and protective devices on machines. Rotary saws are equipped with devices which permit material to come in contact with the blade but prevent the operator's hand from doing so. In some machines the controls which operate dangerous moving parts are located at places distant from these parts. With certain heavy presses the worker after placing material in the press must walk 5 or 10 ft. away to press the button which operates the press. Furthermore, his finger must be constantly on the button; otherwise the press stops.

Secondly, the equipment should not require performance of the operator above his ability. Controls which require unusual forces, speeds of reaction, or patterns of movement obviously are inadequately designed. Displays in the form of gauges and dials which tax the individual's perceptual abilities and are easily misread are equally undesirable. Two examples will illustrate the problems in this area. In certain types of airplanes the controls for the flaps and landing gear are very similar in shape and are located quite close together in the cockpit. In taxiing his craft after a landing the pilot might well pull the wrong lever and hence retract his wheels rather than accomplish the intended operation of retracting his flaps. Situations like this have led to a variety of investigations seeking to develop controls that can be easily distinguished.[13] The ordinary altimeter on aircraft involves the use of three different "hands" on the same dial face, one for thousands of feet, one for hundreds, and one for tens. A common error in the reading of such a dial is 1,000 ft. or more.[6] Obviously an error of this magnitude can lead to very serious results. Other types of dials have been designed which greatly reduce major errors of this kind.

Thirdly, the equipment should give adequate information. The feel of machine controls should be such that the individual can tell when danger is at hand. To some extent the machinist can tell from the feel of the controls when the work in the lathe is getting too tight and is likely

to break and send splinters flying. With highly geared or hydraulically operated controls this feel may be eliminated. Warning signals must be of such a nature as to attract attention and convey a message. The "horn" in aircraft which sounds when the pilot has failed to let down his landing gear upon landing is loud and raucous. Yet after many a belly landing the pilot is pulled out of the plane with the horn still sounding. That which common sense dictates as attention-demanding is not necessarily so. For example it has been found that a steady light is a better warning signal than a flashing one.[6] In the operation of complex apparatus where a check requires the reading of a number of different instruments, difficulty is encountered because different kinds of gauges and dials are utilized. When the signals are arranged so that some easily perceptible pattern emerges errors are not so likely to be made.[6]

One implication of the second principle, the possibility of the individual freeing himself from difficulty, is that controls and releases should be handy. One laundry mangle is so constructed that if an individual catches a hand in the rollers he cannot possibly reach the shut-off control. It is clear that controls should be placed so that they can be easily reached by the operator. The distance an individual can reach has been determined in a number of studies.[41] In general, an individual can reach further to the side than to the front and about equally far up and down. Accuracy of reaching, however, is better forward than to the side even when vision is not a factor.[13] The tendency is to reach too high for targets below the shoulder level, too low for targets above the shoulder level, and too far to the rear for targets located on either side.

A second implication of this principle is that controls and release mechanisms should utilize the best capacities of the individual. For example, the reaction time of the foot usually is quicker than reaction time of the hand.[6] Where either type of emergency control can be used, a foot control would seem to be the more effective. Advantage is taken of this characteristic in many machines where the braking mechanism is a foot pedal rather than a hand lever.

Work Methods in Relation to Safety. As with equipment design, most research on work methods has had the ultimate aim of increasing speed and quality of work. While some studies have been conducted for the specific purpose of increasing safety of performance, they are few in number. Hence little factual information is available with respect to safe work methods.

Work methods are related to safety in three ways.[41] First, though the manner of work is prescribed, the individual may deviate from this procedure and perform an unsafe act. The best way for tightening a screw with a screw driver is to maintain the tool on the same axis as that of the screw. If the screw driver is tilted out of this line the bit is likely

to slip from the slot, and an injury to the hand may result. Similarly, stopping a vehicle short invites rear-end collisions. The consideration here clearly is with skill.

Secondly, incorrect choice of a work method, where several methods are available, may result in an unsafe act. In many situations a given result may be accomplished by several different methods and some of these methods are likely to be safer than others. The decision as to which method will be employed depends upon the judgment of the worker or of his supervisor. For example, in making a small excavation to reach an underground conduit the hole may be V-, U-, or L-shaped. One type of excavation may be better under one set of conditions of terrain and soil, and another type better under other conditions.

Finally, some industrial operations require a check on tools and equipment before they are used in order to avoid unsafe acts. The completion of the operation does not depend upon the inspection, but because there may be breakage or malfunction, the tools and equipment are checked for possible faults before they are used. This preliminary inspection may be as complex as the preflight check given an airplane or as simple as merely noting whether the head and handle of an axe are firmly connected.

By examining each of the above three aspects of work methods in relation to a particular job and by adopting appropriate measures or precautions, safety can be increased. Deviation from the prescribed manner of work could be minimized through training, redesign of tools and equipment, or by changing to a different work method which does not permit extensive deviations. When different methods can be utilized to accomplish the same result, the provision of specific rules to follow in choosing one method over another would eliminate the effects of bad judgment. Where inspection of equipment is necessary, the development of a comprehensive check list and the establishment of procedures for assuring its use are indicated.

Training in Accident Prevention. In accident prevention programs much emphasis has been given to the development of what has been called the *safety habit.* Although there is considerable looseness in the manner in which the term has been used, it is nevertheless amenable to a fairly exact definition. In describing a worker as having or not having the safety habit, reference is usually made to one or more of the following: the degree to which the worker has developed the particular types of skills and techniques of work that reduce the possibility of accidents, the amount of knowledge he has of the operating rules and procedures and of safety regulations, and the extent of his emotional adjustment in so far as it is pertinent to the safety of his performance. The development of these various skills, knowledges, and adjustments, which may

be said to constitute the safety habit, must be given attention in the safety education of the worker.

In the development of adequate skills and techniques of work, one of the most fruitful approaches in the setting up of a safety training program is a study of the types of inadequate worker performances that lead to accidents. An analysis of this sort will point out the significant errors. In training, these errors can be stressed, and special emphasis can be given to their avoidance. As an example of a study of this kind, the accident-causing errors made by four-engine airplane commanders are shown in Table 12-15.[7] The errors given in the table are a simple listing of the errors in performance found in the reports of accident-investigation boards. It is apparent from this analysis that there is no single error of paramount importance, but rather that many different kinds of errors are made. Since some 60 per cent of errors had to do with one aspect or another of airplane control and operation, it could be said that this aspect of performance needed particular attention in training. Not only does such an analysis provide the training personnel with information concerning the major accident-potential areas that need consideration, but it also reveals the specific errors in performance that should be noted. The frequencies with which the various kinds of errors are made serve to point out the particular areas that need intensive emphasis.

In many instances safety training is solely concerned with instruction in operating rules and procedures and in safety regulations. In and of itself such instruction is inadequate since knowledge of safety rules and regulations does not necessarily imply that they will be followed. Nevertheless such knowledge is basic to safe operation. The two factors to be considered in the development of knowledge of these rules are the motivation of the individual to learn them, and the rate at which they are forgotten. It is not enough for the worker to be able to recite a long list of safety rules; rather these rules must be meaningful to him personally. By making the rules really meaningful to him, they are more likely to influence his behavior. Even with proper initial learning, however, the worker can be expected to forget certain of the rules with the passage of time. It follows that periodic tests of knowledge of safety rules should be administered as a check. If the amount forgotten is significant, then retraining is in order.

As was indicated earlier the emotional adjustment of an individual is an important factor in the safety of his performance. The well-adjusted worker can be expected to have fewer accidents than the poorly-adjusted one. Safety education, therefore, should provide for some discussions of the topic of mental hygiene. Orientation of the individual to his job and to the organization, as well as orientation in relation to his social and personal affairs, need treatment in this respect.

Table 12-15. Errors Resulting in Accidents Made by Four-engine Airplane Commanders

Major grouping of errors		Subclassification of errors		Examples of specific errors
%	Type of error	%	Type of error	
28.9	Operation of flight controls	18.7	Landings	Misaligned with runway Landed with one wing too low
		6.0	Approach	Failed to lower flaps Turned into final approach at stalling speed
		2.4	Mid-air collision while in formation	Flew into turning lead ship Overshot in closing formation
		1.8	Take-off	Retracted flaps Allowed ship to yaw and wing to drop
20.5	Power-plant procedure	7.8	Feathering	Tried to correct yaw by feathering propeller Depressed feathering button and circuit breaker simultaneously
		4.2	Power settings	Reduced take-off power too soon Cruised at too high power settings
		3.0	Instruments	Misread fuel gauges Read Autosyn instruments as direct reading
		2.4	Mixture controls	Used emergency-rich too long Set mixture to auto-lean to eliminate torching
		3.0	Miscellaneous	Tried to reduce cylinder-head temperature by opening cowl flaps to maximum Tried to reduce cylinder-head temperature by feathering
12.0	Supervision of crew	4.8	Supervision of crew as a whole	Failed to notify crew to prepare for crash landing Did not post crew member in front of airplane to guide him out of parking area
		3.6	Supervision of copilot	Permitted copilot to stall on landing without intervention Allowed copilot to fly entire pattern with gear down
		2.4	Supervision of navigator	Did not require position reports of navigator Failed to keep navigator informed of fuel consumption
		1.2	Supervision of engineer	Did not require engineer to make continuous oral air-speed readings Directed engineer to make fuel selector changes when not necessary
10.9	Operation of auxiliary controls	7.2	Landing gear, landings	Did not check gear down Intended to lower gear but by mistake retracted flaps
		1.8	Landing gear, take-offs	Retracted gear too soon
		1.8	Miscellaneous	Applied brakes before airplane was airborne Did not use landing lights while taxiing at night at close quarters
9.1	Judgment			Attempted to go around after committed to 2- or 3-engine landing Flew into very turbulent cumulus
8.4	Radio technique and cockpit procedure	3.6	Communication with tower	Did not monitor tower before and during landing Reported wrong ship number to tower
		1.8	Radio equipment	Not proficient with radio compass or command receiver
		1.2	Communication with flight leader	Failed to notify flight leader he was going on 3-engine operation Failed to notify flight leader he was leaving formation
		1.2	Check list	Did not complete check list
		.6	Interphone	Poor interphone procedure
6.6	Observation and orientation			Gave sole attention to instruments, ignoring terrain Failed to observe other airplane in flight
3.6	Safety and air regulations			Violated tower altitude instructions Made instrument letdown through traffic pattern at night

The Clinical Approach to Safety. The clinical method is a process of studying the factors that underlie an individual's behavior. It is essentially a case analysis approach. The aim is to determine the various forces and influences operating upon the individual, and his personality and abilities through which they operate, so that a final integrated picture can be obtained which will be useful in explaining his past behavior and in predicting his future behavior. Both fact finding and analysis are important constituents of the method. Pertinent facts are uncovered from an investigation of the individual's personal, social, and work histories and from various types of descriptions of his emotional make-up and capacities. These are evaluated, weighed, and integrated to form a total pattern descriptive of the person.

The essential element of the clinical process is the inferential and intuitive judgment of the clinician. Controls of the sort that can be effected in scientific investigation are next to impossible. In a scientific experiment, knowledge of the importance of a factor operating in a particular situation can be obtained by holding constant all factors but the one to be studied. For example, the relationship between marital status and industrial accidents among workers can be determined by comparing the safety of two groups of workers who are equated for age, experience, etc., but who differ only in that one group is married and the other is single. In a study of the importance of marital status in a particular individual's accident rate, however, these controls cannot be effected. The determination of what this person's performance would have been had the circumstances been different cannot be accomplished by the scientific method. Therefore the pertinent facts concerning the individual must be evaluated and weighed in the best possible way. The clinical method therefore is necessarily an art; an art in knowing what information to seek, how to obtain it, how to analyze and integrate it into a final description of the person, and how to apply it in suggesting changes in the individual's activities.

From the foregoing discussion it is apparent that the adequacy of the clinical analysis of the accident maker will be a function of the adequacy and ability of the clinician to obtain and interpret information and of the adequacy of the procedures and instruments available to him for obtaining information. It is beyond the scope of the present discussion to describe the qualifications of the clinical psychologist. Suffice it to say that he needs an extensive and specialized professional training, together with experience in the use of diagnostic techniques and analytic and interpretative methods.

The procedures and instruments employed by the clinician in obtaining information about the individual must be evaluated just as the tools used in the selection and classification of workers are evaluated. The

major difference is that the tools used by the clinician should be validated against specific traits or aspects of job performance whereas those used in selection and classification may be evaluated in terms of their predictability of general or over-all performance. For example, if the clinician is interested in ascertaining whether an individual's reaction time is an important factor in his high accident rate, scores on the reaction-time test must be demonstrated to be related to those accidents where speed of reaction is judged to be important. Fletcher, in evaluating the diagnostic value of a test of field of vision, compared the intersectional accidents of a group of automobile drivers with their test scores.[15] He found that, with drivers for whom the test indicated restricted vision on one side, the number of fatal intersectional accidents in which they were involved was six times greater on the deficient side than on the normal side. It is apparent that scores on this test were diagnostic of a particular kind of difficulty of automobile drivers. This type of evaluation is needed for all the procedures and instruments used by the clinician in dealing with the individual case.

The types of information the clinician will require in studying the individual accident case have been considered by Shellow.[36] Information such as the following ordinarily is gathered for the case analysis:

Length of service	Age
Marital status	Health
Home conditions	Interests and background
Job attitude	Personality
Test results	Summary of number and kinds of accidents

Pertinent data will be obtained from a variety of sources. Personnel records can be consulted, reports by instructors and supervisors can be sought, accident records and reports can be consulted, and the clinical psychologist can obtain information directly from the worker through an interview. The data are analyzed and integrated so that a satisfactory explanation of the worker's past unsafe performance is obtained. The clinician can then recommend specific procedures designed to effect an adequate adjustment for the particular individual.

Evidence relative to the effectiveness of clinical procedures is exceedingly gratifying. Shellow reports that the accident rate of a group of high-accident motormen was reduced 82 per cent after having been subjected to clinical procedures.[36] Fletcher has reported a reduction in accidents of some 50 per cent for commercial drivers who had been treated in an accident clinic.[16] He cites more impressive evidence in a study wherein the accident rates of over four hundred automobile drivers were determined before and after they passed through accident clinics and then compared with rates for the same period of a comparable

number of men who were not clinically studied. The clinically examined group showed a reduction of accidents after treatment amounting to more than 60 per cent whereas the accident rate of the untreated group remained unchanged.

The evidence presented above, relative to the effectiveness of the safety clinic as a means of reducing accidents, is only a sample of the positive findings. It appears that the safety clinic is the best single method of reducing accidents in so far as the human factor is concerned. The integration of the work of the clinical psychologist, the social worker, the industrial physician, and the psychiatrist relative to accidents undoubtedly would have a profound effect upon the safety of industrial operations. The major difficulty lies in the fact that the accident clinic requires the service of trained, professional personnel, a service that is expensive. However, in view of the importance of accidents in the physical, mental, and social welfare of workers, the cost of such services should not be the primary concern.

Safety Publicity. No safety program would be complete unless it included some kind of publicity designed to condition the worker toward safer performance. A survey of the literature relative to safety publicity, however, will indicate that the present stage of development of this field is comparable to the stage of commercial advertising some twenty or thirty years ago. The general point of view seems to be that the creation and development of safety publicity is an art and that scientific methods can play little part.

By and large, discussions of safety publicity have been concerned with the effectiveness of different kinds of publicity as they are evaluated by haphazard, casual, and unscientific observation. The effectiveness of systematic market-research techniques in the development of advertising has been adequately demonstrated, and descriptions of these techniques are readily available.[2]

No attempt will be made here to outline in detail the direct application of market-research methods to safety publicity. A few general comments, however, are in order. In the development of safety publicity three distinct areas should be taken into account: first, the obtaining of information relative to the determinants of the behavior in question; secondly, the obtaining of information relative to the reading and listening habits and capacities of the workers under consideration; and, finally, the evaluation of the effectiveness of the publicity.

The first area will provide a systematic attack designed to yield significant information that will be of aid in determining the nature of the publicity, such as the types of appeals most likely to be effective. The second kind of information will be helpful in planning the type of pub-

licity to be used, whether pamphlet, poster, etc. Finally, checks on effectiveness are as necessary here as with any other industrial procedure. Obviously the best check on the effectiveness will be in terms of the reduction of accidents, but partial checks will be provided by tests of the attention-getting and remembering power of the publicity.

REFERENCES

1. Adler, A.: The psychology of repeated accidents in industry, *Am. J. Psychiat.*, **98,** 99–101, 1941.
2. Blankenship, A. B.: "Consumer and Opinion Research," Harper, 1943.
3. Brown, C. W., and E. E. Ghiselli: Factors related to the proficiency of motor coach operators, *J. Appl. Psychol.*, **31,** 477–479, 1947.
4. Brown, C. W., and E. E. Ghiselli: Accident proneness among streetcar motormen and motor coach operators, *J. Appl. Psychol.*, **32,** 20–23, 1948.
5. Chaney, L. W., and H. S. Hanna: The safety movement in the iron and steel industry, *Bus. Labor Statistics*, No. 234, 1918.
6. Chapanis, A., W. R. Garner, and C. T. Morgan: "Applied Experimental Psychology," Wiley, 1949.
7. Crawford, M. P., *et al.:* "Psychological Research on Operational Training in the Continental Air Forces," Army Air Force Aviation Psychology Program Research Report 16, 1947.
8. De Silva, H. R.: "Why We Have Automobile Accidents," Wiley, 1942.
9. Dunbar, F.: "Psycho-somatic Diagnosis," Harper, 1943.
10. Farmer, E., and E. G. Chambers: A study of personal qualities in accident proneness and proficiency, *Ind. Fatigue Research Bd.*, No. 55, 1929.
11. Farmer, E., and E. G. Chambers: A study of accident proneness among motor coach drivers, *Ind. Health Research Bd.*, No. 84, 1939.
12. Farmer, E., E. G. Chambers, and F. J. Kirk: Tests for accident proneness, *Ind. Health Research Bd.*, No. 68, 1933.
13. Fitts, P. M.: "Psychological Research on Equipment Design," Army Air Force Aviation Psychology Program, Report 19, 1947.
14. Fisher, B.: "Mental Causes of Accidents," Houghton Mifflin, 1922.
15. Fletcher, E. D.: Capacity of special tests to measure driving ability, State of California, Division of Drivers' Licenses, mimeographed, undated.
16. Fletcher, E. D.: The effects of special tests on driving ability, State of California, Division of Drivers' Licenses, mimeographed, undated.
17. Ghiselli, E. E., and C. W. Brown: Learning in accident reduction, *J. Appl. Psychol.*, **31,** 580–582, 1947.
18. Ghiselli, E. E., and C. W. Brown: The prediction of accidents of taxicab drivers, *J. Appl. Psychol.*, **33,** 540–546, 1949.
19. Goldmark, J., M. D. Hopkins, and P. S. Florence: Comparison of an eight-hour and a ten-hour plant, *Public Health Bull.*, No. 106, 1920.
20. Greenwood, M., and H. M. Woods: The incidence of industrial accidents upon individuals with specific reference to multiple accidents, *Ind. Fatigue Research Bd.*, No. 4, 1919.
21. Hanman, B.: Placing the handicapped, *Ind. Med.*, **15,** 597–604, 1946.
22. Heinrich, H. W.: "Industrial Accident Prevention," McGraw-Hill, 1941.
23. Henig, M. S.: Intelligence and safety, *J. Educ. Research*, **16,** 81–87, 1927.

24. Hersey, R. B.: Emotional factors in accidents, *Personnel J.*, **15**, 59–65, 1936.
25. Hill, A. B., and G. O. Williams: Investigation of landing accidents in relation to fatigue, Flying Personnel Research Comm. Report 423, 1943.
26. Kraft, M. A., and T. W. Forbes: Evaluating the influences of personal characteristics of the traffic accident experience of transit operators, *Proc. 24th Ann. Meet. Highway Research Bd.*, 278–291, 1944.
27. Minium, E. W.: An experimental study of certain psychological factors in relation to the frequency of accidents in the transportation industry, Ph.D. Thesis, University of California, 1951.
28. Mintz, A., and M. L. Blum: A re-examination of the accident proneness concept, *J. Appl. Psychol.*, **33**, 195–211, 1949.
29. Muscio, B.: A. Two contributions to the study of accident causation. B. On the relation of fatigue and accuracy to speed and deviation of work, *Ind. Fatigue Research Bd.*, No. 19, 1922.
30. National Safety Council: "A Study of Psycho-physiological Selectivity of Employees," Greater Los Angeles Chapter, Fleet Transportation Division, 1946.
31. Newbold, E. M.: A contribution to the study of the human factor in the causation of accidents, *Ind. Fatigue Research Bd.*, No. 34, 1926.
32. Osborne, E. G., and H. M. Vernon: Two contributions to the study of accident causation. A. The influence of temperature and other conditions on the frequency of industrial accidents, *Ind. Fatigue Research Bd.*, No. 19, 1922.
33. Schaefer, V. G.: "Safety Supervision," McGraw-Hill, 1941.
34. Shellow, S. M.: Research in selection of motormen in Milwaukee, *Personnel J.*, **4**, 222–237, 1925.
35. Shellow, S. M.: The accident clinic: how it functions and what it does, *Personnel J.*, **9**, 207–215, 1930.
36. Slocombe, C. S.: How to cut accident costs, *Personnel J.*, **16**, 134–141, 1937.
37. Slocombe, S., and W. V. Bingham: Men who have accidents, *Personnel J.*, **6**, 251–257, 1927.
38. Slocombe, C. S., and E. E. Brakeman: Psychological tests and accident proneness, *Brit. J. Psychol.*, **21**, 29–38, 1930.
39. Swan, E. J.: Economic aspects of social security, *Postwar Economic Studies*, Bd. Governors Fed. Reserve System, **6**, 40–62, 1946.
40. Tiffin, J.: "Industrial Psychology," Prentice-Hall, 1944.
41. U.S. Department of Labor: "Proceedings of the President's Conference on Industrial Safety," 1949.
42. Vernon, H. M.: An investigation of the factors concerned in the causation of industrial accidents, *Health of Munitions Workers Committee, Memo* 21, 1918.
43. Vernon, H. M.: "Industrial Fatigue and Efficiency," Dutton, 1921.
44. Vernon, H. M.: "Accidents and Their Prevention," Macmillan, 1936.
45. Vernon, H. M., T. Bedford, and C. G. Warner: A study of absenteeism in a group of ten collieries, *Ind. Fatigue Research Bd.*, No. 51, 1928.
46. Vernon, H. M., T. Bedford, and C. G. Warner: A study of absenteeism at certain Scottish collieries, *Ind. Health Research Bd.*, No. 62, 1931.
47. Viteles, M. S.: Research in the selection of motormen. II. Methods devised for the Milwaukee Electric Railway and Light Company, *Personnel J.*, **4**, 173–199, 1924.

48. Viteles, M. S.: "Industrial Psychology," Norton, 1932.
49. Wechsler, D.: Tests for taxicab drivers, *J. Personnel Research,* **5,** 24–30, 1926.
50. Wong, W. A., and G. E. Hobbs: Personal factors in industrial accidents—a study of accident proneness in an industrial group, *Ind. Med.,* **18,** 291–294, 1949.

CHAPTER 13

Training in Industry

A progressive industrial organization does not leave the training of its employees to chance but provides systematic instruction on many phases of its operations. The objectives of training programs vary greatly in scope. Some are narrowly conceived, centering the training almost entirely on the goal of immediate increase in production; others are developed on a broader base, assisting the worker not only in his job performance but in other areas of performance directly related to his job.

In the present chapter a case is made for establishing training on a broad base. The worker, as well as his work, is to occupy a central position. Training deals with workers with a variety of backgrounds, some with more ability than their fellows, some with more appropriate work interests and experience, and some with greater zeal to get ahead. The objectives of industrial training should include not only the development of effective work habits and methods of work, but also the development of desirable motives and attitudes. Certainly greater production is a desirable training goal, but solving training problems only in terms of immediate production goals disregards problems in attitudes, motives, and morale, whose solution through training is important and will eventually also bring higher levels of productivity.

TRAINING NEEDS OF THE ORGANIZATION

Who Should Be Trained? A case should be made for stressing the need of a wider coverage of the employed group than is envisaged in most industrial training programs today.[25] Everyone in an organization from the top executive to the least skilled worker is eligible and should be considered a candidate for training. Executive and administrative personnel are undergoing informal training of one kind or another a great deal of the time, most of it voluntarily sought. Many a willing but ineffective skilled worker when given some individual attention shows remarkable improvement. When training includes the continuous development of high-level skills and the improvement of individual motiva-

tion everyone within an organization must be considered eligible and should receive training that will meet his particular needs.

The answer to the question of who is to be trained is directly conditioned by the purposes to be achieved. It should be preceded by the question: For what purpose is training needed? If some particular problem directly related to job performance is concerned, then the individuals eligible for training should be those who have the qualifications in experience, traits, and aptitudes that will enable them, after training, to perform the desired work at an acceptable level. This immediately introduces a problem in selection, for individuals selected for training should have the required experiences, traits, and aptitudes to make the training a profitable undertaking. Some valid standardized procedure should be in force in the selection of the trainees, and as the job to be learned entails more and more individual responsibility, judgment, and initiative, the importance of correct selection procedures increases correspondingly.

Instruction in the Company Culture Pattern. The first step in the adjustment of the new worker is to orient him in the culture pattern in which he is to work. Even before any attempt is made to train him in job procedures and skill, instruction should be given in orienting him to the job, the department, the plant, and the industry.

Particularly is it important to impart to him the company's culture pattern. This is the first step toward getting him to identify with the company so that the goals of the organization will be accepted as goals he should strive for. He should be given facts about the origin and development of the business, the needs filled by the products of the company, the services rendered to the public, the firm's policy relative to quality and price, and other similar facts by which he can learn to identify himself with the organization.

The orientation program should emphasize the responsibilities of the worker as a center of interest. Facts should be presented that will describe the significance of his job and the contribution that it makes in the total organization. Because it is the initial exposure of the worker to a new and unfamiliar working environment it is very important that consideration be given to the worker as an individual within the organization.

Job Training of the Worker. Job training refers to instruction directed to the specific skills, duties, and responsibilities of a particular job. It includes the training of all employees, even the older worker who must undergo preparation for a job unfamiliar to him.

When some particular job-need is to be met such as a vacant position in a job, the available job and worker specifications furnish a basis for determining the areas in which instructional procedures must be prepared.[4,6] Knowing the qualifications required for successful performance

of the job, the individuals eligible for training can be selected. Workers selected for training should have the traits, aptitudes and experience that will enable them to perform the desired work. Only then will maximum returns be received from the training venture.

Job and worker specifications also serve as a basis for determining the correct methods of work to be taught. Many of these methods in present-day industry are standardized, and training officers must see to it that the employees learn accepted standard work methods. This is difficult to accomplish unless the method to be taught is thoroughly impressed upon the instructional staff.[29]

Training for Promotion to Higher-level Jobs. A common method of filling the top jobs in a unit such as an office, department, or plant is to advance those who are already working in the unit. One of the deficiencies of this procedure is found in the original selection of the workers. If at the time of selection no attention was paid to the potentiality for growth of the selectees, then the workers available in a unit may not have the aptitude to perform the top job at a really high level of proficiency.

The question of training for advancement resolves itself into planning instruction to ready a given worker for the new responsibilities he will exercise. In foremen and supervisory jobs this requires training in social relations including instruction in understanding the worker's need for personal recognition. One of the traits which differentiates the good from the poor supervisor is an ability to get workers to want to be proficient producers. Training in interpersonal and leadership relationships is of assistance to the potential supervisor in developing and expressing this ability.

In connection with this type of training an important task of management is to arrange a scheme of stepwise promotional advancements within the company's organizational structure. Employees should then be informed concerning the kinds of experience and knowledge a worker is presumed to obtain in preparation for any given promotional step. To have such a promotional scheme to "shoot at" is of great value in stimulating the worker to perform effectively in any relevant training program that is organized.

TRAINING NEEDS OF THE WORKER

Improving Work Habits. Sometimes a worker gets the idea that he ought to improve his performance on the job. This idea may have come about because his usual level of production came out second best when he compared it with that of a fellow worker, with his own performance of an earlier date, or with some arbitrary standard of performance set by

himself or by others. There are several reasons why his performance might be thought improvable. He may have adopted faulty response patterns when he first learned the job. Since starting the job he may have unknowningly changed his response patterns for the worse. The nature of the work may have changed, and he may have failed to adjust to the changes. He may have discontinued supervised learning before reaching his potential maximum performance and never have received the assistance necessary to develop more proficient methods. What is important is that the worker seems willing to attempt a readjustment of his response patterns in order to improve his proficiency. He needs aid in unlearning his less effective methods and in mastering more effective ones. He needs opportunity, under guidance, to effect this readjustment.

Three conditions must be met if instruction is to be successful. First, there should be some gain or reward for the worker to seek. This need not be an increase in income, but it must be in a form which the worker considers of worth to himself. Secondly, a study of the worker's performance should be made to determine areas wherein improvement is possible. This may involve a job and worker analysis. Thirdly, readjustive responses should be planned which will come within the psychophysiological limits of tolerance of the worker and which at the same time will effect an improvement in his job performance. The three ingredients of successful human learning are then present, viz., knowledge that a change is needed, knowledge of how the change can be made, and desire to make the necessary readjustment.

Training in the Formation of Goals. Training in establishing goals may be useful to many workers. Such training should be proffered to the person who complains that he is in a "rut." It is a form of training applicable to the problems of the bright but dronish worker. It might actually be sought by the individual who repeatedly remarks: "I wish I knew what I wanted to do." In each of these cases there is a need for defining and accepting tangible goals and establishing workable plans for making a start toward these goals. The goals need not necessarily be directly connected with the job. They can be cultural in nature or can be changes in the worker's personal habits. What is common is the need for instruction in initiative and drive. The problem is one in motivation, not one in work methods or skills.

In this type of training the function of the training officer is a very important one, calling for considerable tact and understanding. The instruction first consists in getting the worker to formulate a goal. Often the sign that the individual is in need of training comes in the form of a complaint about his own progress or his future prospects for success. Such a sign is not to be interpreted to mean that the worker is convinced that he needs a change. It may merely indicate that he might be open to

suggestions. It is the task of the training officer to stimulate the worker to feel a personal need for setting a goal and to get him to formulate a goal which he will identify as his own. When this is accomplished the training officer can then assist the worker to lay plans for readjusting his response patterns in the direction of attaining the goal.

The goal should have certain characteristics. It must come readily within the capacities of the worker to achieve—it cannot then be too idealistic or too remote in time. If the goal itself cannot be realized in a reasonably short time, then it should contain subgoals which can be more readily attained. The goal should be sufficiently concrete in nature that the worker will be able to note progress toward its realization in some kind of tangible results accomplished. The goal must be worthwhile to the worker, and its attainment a significant contribution to his future development. If the goal does not have a strong pulling power the worker is likely not to succeed in the necessary readjustments which the realization of the goal demands.

Instruction toward Better Job Adjustment. This problem concerns the worker's failure to realize his established job goals. This failure may be due to one or more of several causes. A worker may be a vocational misfit because of circumstances not under his control, that is, for economic reasons he may have had to discontinue preparation for one job and accept work in another. A worker's ignorance of what job opportunities were available to him when he came to the company may have led him to accept his present job. A worker may have made a wise choice of a future career but at critical points lacked the drive to get over the tough spots and thus had to postpone temporarily further job training. A worker's choice of a job may not have been a correct one in terms of his aptitudes. A worker may be doing the present job in the hope that a break will come which will allow him to get a job in a related area where he believes he should be. It is apparent that being unhappy on the present job may stem from many different causes connected with the worker's failure to advance toward his desired vocational goal.

The kind of instruction to be offered is to be adapted to the particular problem presented by the worker. Sometimes vocational information that will steer the worker into a self-study program or correct his erroneous ideas about goals that are futile is all that is required. Sometimes the instruction must be directed toward strengthening the desired goal and of giving the encouragement necessary to arouse a drive to enter a special training program. Sometimes merely a transfer of jobs accompanied with refresher training is sufficient. In all efforts to help the worker the underlying philosophy is that "you *can* teach an old dog new tricks" when there is adequate motivation and appropriate instructional facilities.

In the main this type of training is an attack upon the problems of classification, and the resultant gain to the employer is usually considerable. The objective is to assist the worker to develop insight into his particular aptitudes, and to provide within the company framework the opportunities through which his talents can best be utilized and he can realize maximum self-expression. With the attainment of these objectives, worker discontent will disappear. With compatibility established between the requirements of the job and the worker's abilities, interests, and temperament, there should follow a high level of proficiency over a long period of time.

CHANGES IN BEHAVIOR DURING LEARNING AND TRAINING

Areas in Which Improvement from Instruction Can Be Expected. Improvement to be anticipated from adequately conducted instructional programs can be looked for in almost every area associated with worker performance. Regardless of what measure is used for evaluating the training, an increase in effectiveness is usually found. A review of the results of training programs in industry revealed the following: increased worker satisfaction, higher levels of morale, increased interest in the work, closer cooperation between management and the workers, reduction in sick time, reduction in absenteeism, improved safety records, increased level of output, reduced variation in output, and decreased labor turnover.

A typical example, reporting the improvement in skill resulting from a planned, intensive training program, is found in a study by Langley and Edwards.[29] These investigators made comparisons between formally trained and untrained groups, a type of scientific approach all too infrequent in studies of industrial training programs. They report that the group receiving formal training reached a level of proficiency in 14 weeks that was attained by the informally trained workers only after a period of 3 years. In a task primarily involving hand and finger movements they found that less than 15 per cent of the trainees failed to attain the high standard they had set as a goal for all workers. Even these slower workers showed significant improvement, and their failure to attain the goal was attributed not to lack of ability but to lack of sufficient motivation. In addition to an increase in hourly output of about 50 per cent, the authors report that there was no resentment on the part of the workers in being trained, that absenteeism almost vanished, that labor costs were materially reduced, that labor turnover was greatly reduced, and that the operators directly benefited from an increase in earnings.

Changes Occurring in the Behavior Being Learned. In a typical learning situation the rate of improvement is initially quite rapid and then gradually becomes less and less until the learner approaches a final level in performance called a terminal plateau. Figure 13-1 presents a typical learning curve and represents the improvement in performance of new workers in a printing plant.[24] It will be seen that the final level of proficiency was not achieved until after some six months of practice. For many jobs, gains are registered over a much longer period than that represented for the printing-plant workers.

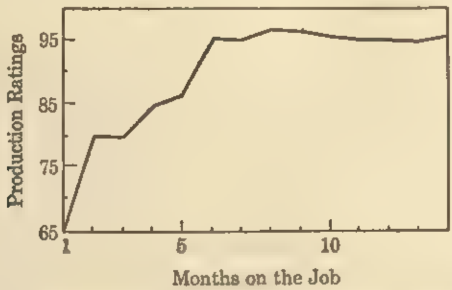


FIG. 13-1. Improvement of new workers in a printing plant.

Changes from practice are to be looked for in performance areas other than increased rate or amount of production. Through practice certain phases of the action patterns are automatized, superfluous activities are dropped out, precision of movement and smoothness of coordination are effected, excess tension is eliminated, the energy requirements of the total performance are greatly reduced, and the trainee develops self-assurance and confidence in his performance. The result is the attainment of an effective job performance within the psychophysiological limits of tolerance of the worker.

Extending the Learning Time to Attain a High Level of Proficiency. Training courses in industry are frequently of short duration, ranging from a few days to a few weeks. Sometimes, of course, there is little need of formal training and the worker can be placed on the job after a few practice sessions. From known empirical findings, however, it would appear that much industrial training is of insufficient length to capitalize on the full extent of the period of high gains known to characterize the development of complex behavior. Three points will be discussed that throw doubt on the notion that training need only last a short period of time.

First, the presumption is made that the initial rapid gains are accomplished in a few days or weeks. While many simple tasks can be learned in a very short period of time, the complexity of most jobs is such that

the initial period of high gains is spread over a fairly extensive period of time. As previously shown in Fig. 13-1 workers in a printing plant did not reach the final level of proficiency until after some 6 months of practice. For other jobs the learning period is even longer.

A second consideration concerns the need for training the worker beyond possible intermediate plateaus. A typical example is given in Fig.

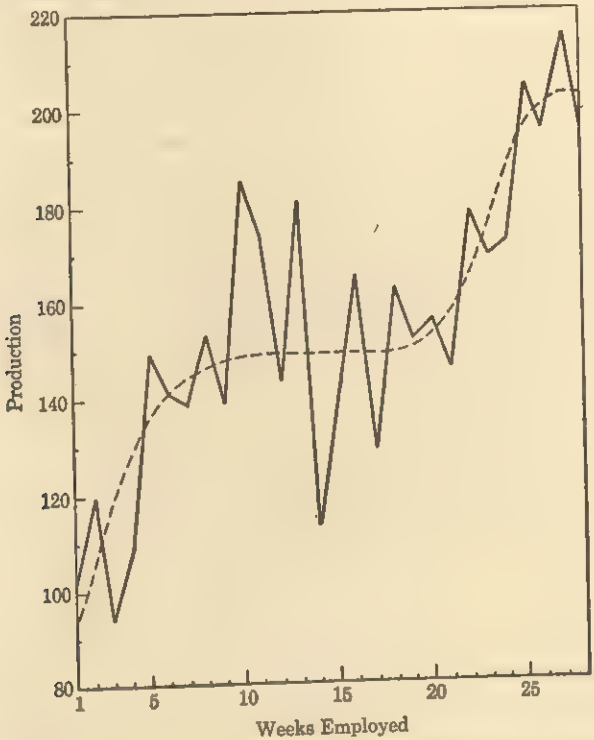


Fig. 13-2. Production of a new taxicab driver during the first 28 weeks of employment.

13-2 which shows the course of improvement in production of a new taxicab driver. It will be noted that an intermediate plateau occurred around the seventh week at which time his performance was 50 per cent greater than his initial performance. This plateau, however, proved to be only temporary and after some 12 additional weeks of practice his performance again began to improve with the indication that it would level off at about 100 per cent better than his initial performance. If an estimate of the operator's final level of achievement had been made during the intermediate plateau, it would have been approximately 50 per cent less than he actually showed himself capable of producing. Mistaking the

intermediate plateau for the final plateau would have led to the conclusion that learning had been completed long before the worker actually had attained his maximum performance.

Intermediate plateaus are ascribed to two different sets of factors, viz., changes in motivation, and learning new and better ways of performing the same task.[5] As the new worker approaches his task initially he does so with some measure of enthusiasm. However, as he examines his progress he may find it unacceptable, especially as he compares his performance with that of the more experienced persons with whom he works. His performance may seem quite inadequate to him, and he may become discouraged. This lowered motivation will result in further progress being arrested. It is only after his interest and motivation are again aroused that further improvement in performance is manifest.

In many instances intermediate plateaus arise from the fact that the individual has reached his psychophysiological limit in the performance of a task. After a time, however, he may discover a new method of work and then will show a further improvement in performance. Learning to use a typewriter is a typical case. At first the trainee types individual letters, spelling out each word as he types. He soon reaches maximum speed and no further gains accrue from practice. After a while he attacks the problem differently, thinking in terms of words rather than individual letters as he types. The result is that he is really learning a new task, and with this better method of work his performance again improves. After a second intermediate plateau he may learn to attack the problem by thinking in terms of phrases rather than words and show a still further gain.

A third consideration that throws doubt on the notion that training can be accomplished in a short time arises from the facts known about overlearning. As soon as a trainee has reached his final level of performance he is considered to have been adequately trained. Empirical studies show, however, that additional practice—called overlearning—is very beneficial. The greater the amount of overlearning the better established will be the newly learned response-patterns and the lower the probability that forgetting will occur or that the worker will regress to former ineffective response-patterns when he is confronted with a difficult situation.[28] Training should provide the opportunity for the worker to practice under supervision beyond the point where he has just barely achieved maximum performance. The usual rule-of-thumb is to allow as much practice for overlearning as was required for the original learning. The benefits to be derived from different amounts of overlearning vary with the nature of the work, and hence the optimal amount should be determined for each particular job.

INDIVIDUAL DIFFERENCES IN RELATION TO TRAINING

At first glance it would seem unnecessary to discuss and emphasize the importance of individual differences among workers in relation to industrial training. If there has been one finding in psychology established beyond reasonable doubt, it is the fact of individual differences. Not only that there are differences between individuals in capacity to profit from training, but also that these differences persist to some extent despite efforts to erase them.

Extent and Amount of Individual Differences. The extent of individual differences is greater than directors of industrial training are usually willing to admit. Differences are to be found in every form of behavior. Although experimental studies have been preponderantly centered on abilities, aptitudes, and skills, differences just as pronounced are to be found among workers in drives, motives, interests, aversions, social behavior, personality traits, attitudes, opinions, and other modes of behavior.

In regard to any estimate of the amount of individual differences in work performance, the nature of the group being studied must be taken into consideration. Obviously, the more heterogeneous the sample of workers, the greater the probability of getting large variations in performance. For our present purposes run-of-the-mill groups of workers can be considered as possessing the differences with which training officers will be faced. When such groups are studied, and the amount of the differences between the individuals is measured, it is usually found that the best worker is some two to eight times better than the worst worker. With differences of these magnitudes present in most work-performance situations, some thought certainly should be given to such differences when organizing industrial training programs.

Individual Differences in Learning. Variations in learning performance among different workers take several forms. First there are differences in initial performance. In addition to variations in native learning ability, these differences may be attributable in part to differences in motivation such as willingness to embark on a training program, differences in ability to understand what is to be learned, and differences in ability to follow directions. Obviously there are more factors at work than simply variations in the aptitude or ability to learn.

A second difference is in general rate of improvement. Some workers learn more slowly than others. It is not to be inferred from their slower rate of improvement that they will never be as proficient on the job as their fellows who learn more rapidly. Nor is the opposite interpretation true —although frequently encountered—which holds that the slower learner gets a better grounding and therefore will eventually become more pro-

ficient than the worker who learns more rapidly. Neither of these generalizations can be made of all learning, but facts supporting each of them can be found in different kinds of specific learning situations. One fact does stand out; there are differences in learning rate among individuals in all types of learning situations.

Another characteristic in which individual differences are found is the ultimate level of proficiency attained at the completion of training or at any other later time after the worker has had some practice on the job. It is variation in this characteristic that training officers have labored diligently to erase. Experiments indicate that continued practice does not remove all differences among learners.[1] There is an ultimate limit set on the rate of learning by the psychophysiological mechanisms involved in any given task to be learned, and this limit is not the same for all learners. This is true whether the activity is one of coordination between muscular reactions or one primarily mental in nature such as reasoning, remembering, etc. Furthermore, psychophysiological limits exist for various characteristics of behavior such as the rate of movement, the preciseness or fineness of the coordination, and the ultimate level of difficulty at which acceptable performance can be achieved.

The relationships among the three variables—initial performance, rate of learning, and final level of proficiency—will vary considerably with different tasks. An example of the possible relationships to be found is furnished in an investigation by Blankenship and Taylor, who studied the performance of power-sewing-machine operators on three different jobs for the first 50 weeks of work.[7] Initial performance was measured by success during the first 10 weeks, speed of learning by the time required to meet a certain standard of proficiency, and ultimate proficiency by the production for the entire period. The results, given in Table 13-1, show rather marked differences among the jobs but in general indicate

Table 13-1. The Relationships between Initial Performance, Speed of Learning, and Ultimate Proficiency in Learning Found for Three Power-sewing-machine Jobs

Correlation between—	Type of job		
	Trimming	Covering	Hemming
Initial performance and ultimate proficiency	.68	.01	.66
Speed of learning and ultimate proficiency	.74	.52	.59
Initial performance and speed of learning	.66	.41	.32

that ultimate proficiency is more closely related to speed of learning than to initial performance. The relationship between initial performance and speed of learning is fairly high.

Differences Due to Instructional Methods. Other factors determining individual differences in rate and quality of learning are related to the appropriateness of the training methods utilized. Some of the factors to be considered in evolving the most appropriate procedure for a given training situation will be discussed later. It is sufficient here to recall the common occurrence in industrial training of variations in the learning rate of new employees under different foremen-trainers. The approach and method of some foremen get much faster results from the learners than those characteristic of other foremen, indicating that differences in methods of teaching produce differences in performance of the trainees.

Emphasis upon Individualized Training. The presence of individual differences in the characteristics described in the preceding paragraphs makes it necessary for training officers to adjust training programs to individual needs. When the activities to be taught are simple in nature and the trainees have been selected carefully, individual differences present little difficulty to the training officer. The more complex the behavior to be developed and the less carefully the trainees are selected, the more important are the problems arising from individual differences.

Differences in the purposes of training programs are associated with the emphasis to be placed on individual differences. When training specifically concerns knowledge and skill on the job and requires an intensive development of the worker in restricted fields, the emphasis may be somewhat different than when training involves problems concerned with the worker's adjustments in other than the immediate job situation. Within the latter, more extensive type of training there must develop a more individualized approach—a concern with the problems of the worker as an individual. In many instances the actual training given may be one of group instruction, but certainly the isolation of the problems needing attention, and much of the training necessary in the solution of these problems, will require that the worker be dealt with as an individual.

What is being suggested is that individual differences in capacities, motivations, and experiences be given more recognition. Taking such differences into account will not only result in better training methods but also in a clarification of training objectives.

Adjusting Instruction to Worker's Abilities. Adjusting the instruction to the worker's abilities has many aspects including the following: starting the instruction at a level of difficulty commensurate with the worker's readiness, relating the new training experiences with experiences familiar

to the worker, and adapting the rate of instruction to the quality and rate of the learning being manifested by the worker. Obviously, the significance of each of these factors in reference to the worker's readjustment will vary with the nature of the training problem and to some extent with the kind of instructional procedure that can be applied.

Training should be adjusted to the worker's readiness to profit from it. The worker may fail to learn the necessary readjustments or be greatly retarded in his progress if he starts training before he is prepared. Readiness to begin learning includes factors of motivation, interest and desire, as well as factors of aptitude, ability, and experience. The time at which the worker is ready will depend upon the problem, the nature of the readjustment to be accomplished, and the type of instruction that can be used. The point to emphasize is that maximum returns from training cannot be realized whenever any of the factors of readiness exert a negative effect on the worker's progress.

Instruction should adjust the new tasks to be learned to the particular background of each trainee. The meanings of the new tasks can be assimilated more easily when they are directly related to familiar meanings, when they can be applied in the present response patterns of the worker, and when the worker understands the contribution the new learning will make to his future status. Meshing these new experiences with the experiences of the past will receive further consideration later.

Rate of improvement should be adapted to the psychophysiological limits of the individual worker. To force every worker to progress at the same rate is to disregard the principle of individual differences. It is to exert undue pressure upon the very slow learner, and to champion mediocrity by requiring less progress from the more capable worker than he can readily achieve.

ASSIMILATION THROUGH COMPREHENSION

Assimilating the New through the Old. The comprehension and development of new things—whether they are information in the form of facts, personal habits in the form of incentives to excel, or skill in the form of motor coordinations—can be accomplished only through the knowledge, experience, and abilities which the worker has already developed. To a large extent the rapidity with which the assimilation of the new learning will occur is dependent upon how closely the individual's past is related to it and upon the degree to which his past can be directly focused on the modifications of behavior that are to be accomplished.

The direct application of the old to the new cannot be left to chance or be allowed to develop incidentally to the instructional program. A purposeful and conscientious attempt should be made by the instructor to

relate that which is to be learned to some item of knowledge, experience, or ability familiar to the worker. This will be seen as no mean task when it is realized that the relations of the new to the old will vary greatly for different workers. Of course, if the individual is left to his own devices he may discover the relevant relationships and principles. Postman and Jarrett have shown, however, that the performances of those who are directly taught these relationships and principles are superior to the performances of those who discover them without help.[34]

Logical Compared with Rote Learning. The best method for effecting a strong relationship between the old and the new is one that stimulates comprehension and understanding on the part of the learner rather than one which primarily advocates mere repetition.[31] This is the old argument between logical learning and rote learning. In rote learning the task is repeated in a drill fashion as in learning the multiplication tables. In logical learning an attempt is made to introduce as many meanings into the new experience as is possible, and through these meanings to assimilate the new into the past experiences of the worker.

Cox made a study of these two teaching methods in training workers to assemble electric-lamp holders.[11] He compared a mechanical repetition of the task with instructed training in which general principles and meanings underlying the operations were explained. The group using the logical learning method showed significantly greater gains in proficiency than the group using the rote method. From many similar studies it is clear that the more meaningful the readjustment appears to the learner the more rapidly it will be acquired.[27]

Learning General Principles during Training. One objective toward which the new meanings should be directed is the establishment of general principles of attack that will encourage the trainee to modify his behavior when difficulties are encountered. A common finding in studies of learning is the persistence that the trainee shows in repeating responses that have proved inadequate. This is particularly true in problems having no ready solution. The learner needs to be taught to vary his attack and not perseverate by repeating an inadequate response. This can be done through instruction in logical learning. The reduction in perseveration that follows greatly facilitates the learning.

This principle is illustrated in a study by Maier in which the instructions on general methods of attacking a problem were varied with two groups of subjects.[30] Before beginning work one group was merely told what the problem was and given the equipment required to solve it. The other group received a short lecture on general methods of attacking problems. The lecture was not directed specifically to the problem at hand but was directed to the need of avoiding perseveration of inade-

quate reactions. A significantly greater proportion of the second group solved the problem within the time limit set.

Effect of Comprehension on Motivation to Learn. Comprehension of the new task will make the learning more rapid and the retention more stable, but the effect does not end here—there will be a direct facilitative effect upon the motivation of the learner. Everyone has had the experience of wanting to give up the learning of a task because it proved too difficult to do or to understand. Wanting to quit or wanting to go on is a matter of motivation. It requires a higher level of motivation to learn in the face of difficulties or lack of understanding. With comprehension and insight there follows more interest and desire to learn with the result that the learning becomes easier.

MOTIVATIONAL FACTORS IN LEARNING

The motives for learning are largely the same motives as those actuating the worker on the job. He desires to learn because he is in need of a job, or he wants to qualify for a higher position, or he expects increased financial returns, or he anticipates greater prestige, or because of other results that he thinks will further his personal status. However, there are certain problems of motivation of particular importance to the training situation and these will be dealt with in the following topics.

Personalized Motivation. One of the primary requirements of individualized training is that there be personalized motivation. In personalized motivation each worker must have goals toward which training will advance him, and in the absence of such goals the first function of training is to assist him to establish them. Training in motivation, in goal-setting, and in foresight and planning should not be ignored or made incidental to training in skill. What is needed is planned instruction that will assist the worker in identifying himself with goals that will result in further advancement in remuneration, satisfaction, and security.

The first step in obtaining personalized motivation is to create in the worker a desire to undertake the training. This can be done by explaining in detail the conditions that give rise to advancement and promotion in the organization and how the instructional programs will enable him to accelerate his progress. He should be given facts which support the value of any training to which he can become identified, and be informed on how he can contribute to the planning of such training. He should be made acquainted with the returns he can anticipate from the training, and the demands that it will make on him. Until the worker has become convinced of the need of training, of the worthwhileness of the returns, and of the work entailed, the level of motivation will be low, and learning will be perfunctory and slow.

Establishing Attainable Goals during Training. A marked discrepancy during training between a worker's level of aspiration and his level of achievement can have a traumatic effect, arousing feelings of inadequacy and failure, causing him to become frustrated, and producing a plateau in his learning. Coch and French have shown that these adverse effects arise from frustrating situations,[10] and may lead to resentment toward management and an increase in labor turnover among new workers.

The solution to this problem is to provide reasonable goals during learning—goals that are attainable within the ability of the worker. Trainees should be instructed in the amount of progress that they can expect to make in any readjustive learning they attempt. This amount should be a realistic estimate based on the performance of former trainees and not on the performance of experienced workers. In the early stages of learning, the goals should be set at low levels of accomplishment, care being taken to assure the worker that the goals are attainable. The level of achievement can thus be made to correspond with the level of aspiration, and feelings of inadequacy, failure, and resentment will be minimized and frustration avoided. This type of approach has been used with considerable success.[10]

Knowledge of Results. It is difficult to see how an individual can be expected to learn rapidly if he has no knowledge concerning his achievement, yet in many training situations the worker is in exactly this situation. Consider, for example, the position of a person newly placed on an electrical-parts-assembly job. After being shown how the job is to be done, the worker is left to his own devices to practice and to learn. Undoubtedly the task can be done in a variety of ways, *e.g.*, applying a small or a large bit of solder to fasten a wire to a terminal. If the final assemblies are inspected by another worker who corrects inadequate work but who does not inform the assembler of errors, the assembler will never know when his work is satisfactory or when it is unsatisfactory.

Knowledge of learning achievements not only helps the individual by indicating which methods of work and approaches to the task are effective and which ineffective, but is tied in with motivation by indicating to him his level of achievement. A number of investigations have shown that the rate and quality of the learning are conditioned by the extent to which the trainee has knowledge of his achievements.[16]

Praise and Reproof. A further problem that arises concerns the relative emphasis that should be placed on success and failure during the learning. The question can be posed as: Is it better to praise the trainee and therefore reward him for any improvement he achieves or is it better to reprove him and thereby punish him for the errors that he makes? The results of studies of these two factors indicate that both are effective.[31] This is not unexpected since it is important for the trainee not only to

know when he is doing something correctly but also to know when he is doing it incorrectly. By this means he will know which responses, methods of work, and approaches are correct and which are wrong. If he is given information of only one kind he can judge the other only by implication.

Investigations of learning suggest that stressing praise is slightly more effective than stressing reproof.[31] Attitudes developed during training carry over to other phases of job performance. If the learner is praised rather than reproved, it would appear that at the very beginning a foundation is developed for more favorable attitudes on the part of the worker toward management and the organization as a whole.

Competition and Rivalry. In some organizations there is a practice of fostering a spirit of rivalry among trainees in order to increase motivation to learn. In effecting this spirit special awards are given to those who show the greatest progress and demerits of some kind to those whose progress is least. Sometimes this is accomplished by giving bonuses or other rewards, or by requiring extra practice time. The same ends can be reached by public declarations such as posting a ranking of the trainees on a bulletin board.

As indicated earlier, the development of a sense of failure is not conducive to rapid learning. Singling out certain trainees from others through some form of demerits would appear to do just this. But by being put in a competitive situation the individual is encouraged to adopt any means he can to better his position. It is not unexpected to find that while competition does in fact facilitate improvement during learning when this is measured by amount of work, these gains are often achieved at the expense of quality of work.[14]

THE FUNCTION OF PRACTICE

Mere understanding of a new form of behavior is seldom sufficient in itself to enable the worker to perform it effectively. Practice is necessary to perfect any response and to maintain it at a high level of proficiency.

The Meaning of Practice. Practice refers to the repetition of the task with the objective of improving performance. There is no standard or criterion which can be applied to determine when a task has been practiced sufficiently. An activity may be considered learned when there has been very little practice. In this instance the criterion usually will not require a highly proficient performance. For a large proportion of industrial tasks considerable amounts of practice are required to reach an acceptable level of effectiveness in performance.

Habituation. The criterion of learning may be set very high, and one of the functions of practice is to develop the task to the point at which

it can be performed habitually. In the beginning of learning, whether the task consists primarily of mental or of motor activity, its performance requires conscious direction on the part of the learner. For example, in the operation of a lathe the worker has to respond with both hands at varying speeds of movement and in various positions in space. These actions are dependent upon spatial judgments, which in turn are dependent upon visual acuity and sensitivity. Performance at the beginning of training is slow; the movements awkward and uncoordinated. Through practice these slow, uncoordinated movements of the beginner are displaced by the rapid, precise, finely integrated movements of the expert. Eventually a large portion of the sequence of response may require little or no attention on the part of the worker. This habituation frees conscious direction so it can be focused upon the more difficult phases.[8]

In formal training situations wherein specific job responses are undergoing improvement, the problem arises of determining how much of the habituation should be accomplished in the formal instructional program and how much should be left to be achieved on the job itself. No single answer can be given. Such factors as the following must be taken into consideration in arriving at an answer: the nature of the behavior to be learned, the stage of development to be attained, the amount of difference between the trainee's execution and that of experienced workers alongside of whom he will be working, and the kind and amount of tutoring he will receive from his supervisor and fellow workers.

Reduction in Feelings of Effort and in Energy Expended. Coincident with the reduction of conscious direction occurring in the learning of a task, there is a reduction in the mental and physical effort required to perform it. Considerable mental effort is involved in the acts of attention and concentration. In the beginning of training so many aspects of the act to be mastered are competing for attention that the learner cannot afford to relax at all but must continually maintain a high level of vigilance. As each aspect is reduced to a habitual level less effort is required to perform it. More work can then be accomplished before feelings of tiredness set in.

During learning, the physical energy required to perform the act will be lessened through a reduction in muscular tension. In the initial stages of learning the worker is tense, knitting his brows, gritting his teeth, squinting his eyes, and in other ways expending energy not directly part of the task. Much of this tension arises from fear of making errors and from awkwardness and incoordination in the muscle groups which are essential to the task but over which precise control has yet to be achieved. As practice continues this excess tension declines, and in a highly perfected reaction may disappear entirely.[21]

Changes in the Form and Timing of Component Motions. Practice not only increases the effectiveness of performance but results in changes in the timing of motions and in the final patterning of motions which is eventually adopted and used. Complex performances always involve the coordination of several component activities. The organization of these part activities is often one of arranging temporal linkages between individual reactions. This may be illustrated by an experiment in which a simple motor task was studied. The individual was required to tap six keys in the order 1, 2, 3, 4, 5, 6, 1. The time necessary to make the movements between each pair of keys was measured. In Table 13-2 the time

Table 13-2. The Relative Speed of Movement at the Beginning and End of Learning in a Key-tapping Task

Movement between keys	Relative time to make the movement		Ratio of late to early trials
	Early trials, %	Late trials, %	
1 and 2	13.9	16.6	119
2 and 3	12.2	11.6	95
3 and 4	15.0	16.2	108
4 and 3	14.3	13.1	92
3 and 5	11.8	12.4	105
5 and 6	16.4	16.0	96
6 and 1	16.4	14.1	86
Total	100.0	100.0	

for each movement is expressed as a percentage of the total time to complete the sequence. Although at the end of learning the total time for the sequence decreased about 20 per cent, it will be noted that some of the movements required relatively more time in later trials than in earlier ones, and others required relatively less time. Thus the relative amounts of time of the component movements in the final perfected act were not the same as those found in the initial phases of the performance.

During practice, changes occur in the patterning of the component responses. Reactions in various muscle groups are integrated and coördinated as parts of the larger unitary act. As previously pointed out, in the beginning of learning these individual reactions are crude and awkward. Fineness and precision of movement are learned through prac-

tice. In this process some individual reactions may be discarded and others added. New linkages are found and old linkages dropped out. Short cuts are discovered and introduced into the sequence. Variability of the movements is reduced and smoothness of coordination gradually acquired.[8]

Instructional Guidance in Relation to Practice. Over and above practice, the worker needs to be guided with respect to specific methods of

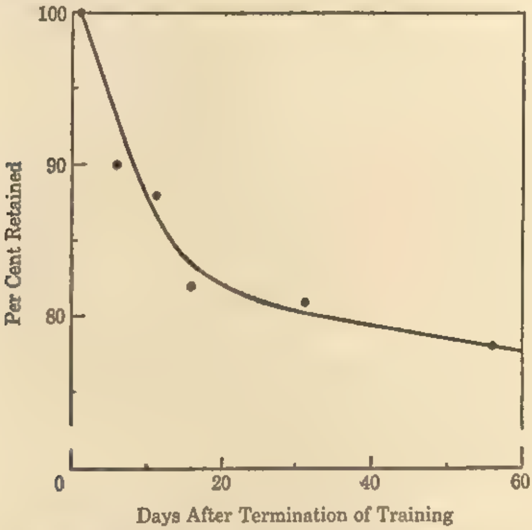


FIG. 13-3. Retention of material learned in a basic electronics course.

work. If he is not given this guidance he may adopt ineffective methods. When guidance in effective methods is given, improvement is more rapid and higher levels of performance are attained. On purely logical grounds it would appear that the effectiveness of guidance should be greater during the earlier rather than the later stages of learning. Studies of guidance during learning substantiate this point of view.[9] The beneficial effects of guidance diminish as the guidance is introduced later in the practice period. Excessive amounts of guidance may even hinder learning since the worker has insufficient opportunity to perform the act on his own, freed from the "crutches" supplied through guidance.

The Effects of Lack of Practice. The general shape of the curve of forgetting has been empirically established. After cessation of practice or use the worker's performance becomes less effective as time passes. Greater losses usually occur in stages immediately following learning than in later stages so the curve of forgetting often approximates the form presented in Fig. 13-3.[39]

Losses in performance are not simply due to physiological deterioration or decay incident to disuse. Rather these losses are a function of many factors.[31] Thus the kind of activity in which the worker engages between the initial learning and the reproduction influences the accuracy of the later performance.

With mental-type tasks lack of practice may result in distortion in that which is later reproduced.[2] For example, if a complex procedure that has recently been learned has not been practiced for a considerable time, not only will the worker be unable to reproduce certain phases, but many of those he does reproduce may be performed incorrectly.

With skilled motor activities, particularly those that have become habituated by overlearning, the loss through forgetting may be very small.[31] For example, the loss in accuracy in aerial gunnery of highly trained fighter pilots who were inactive for about seven months was found to amount to only 6 per cent.[13] This finding should not be construed to mean that the amount of forgetting in motor tasks is invariably small. Such small losses occur only when the individual has had a very large amount of practice, that is, when there has been a great deal of overlearning.

Practice as Retraining. Practice is not to be considered important in the learning solely of the novice or the new employee. Many studies have demonstrated that so-called "experienced" workers can effect improvement through systematic practice. An investigation carried out during World War II with experienced fighter-pilot instructors can be cited as an example.[13] These men, highly accomplished in aerial gunnery, improved 450 per cent in accuracy of firing after 50 practice missions. Improvements from retraining have also been found in the industrial situation.[36] Langley and Edwards demonstrated the feasibility of improving the proficiency of acknowledged "inefficient" workers.[29] As with new learners, improvement of experienced personnel may be manifested in speed of performance, accuracy of performance, ease of performance, or change in the method or form of executing the task.

The question of when retraining should be instituted is not easy to answer. In part it will be determined by the amount of time that has elapsed since the original training and in part by the opportunity the individual has had to practice the activities since training. It is clear, however, that the long-term benefits from retraining will be greater the earlier the retraining can occur after the original learning.[37] Retraining will then be introduced when the effects from forgetting are at a minimum.

TRANSFER FROM THE TRAINING SITUATION TO THE JOB SITUATION

The problem of transfer of training is concerned with the degree to which one subject or skill influences the subsequent learning or performance of other subjects or skills. The question can be asked: To what extent will the movements, methods of work, knowledge, attitudes, and insights learned in one situation affect performance in some other situation? In a larger sense all training is concerned with transfer, since that which is learned during training is expected to be utilized—that is, transferred—to the actual job situation. In a narrower sense, transfer of training is involved in particular training methods and procedures. Thus the utilization of visual aids, synthetic trainers, etc., to develop skills and knowledges, presumes that what is learned through these means will carry over and facilitate learning of the actual job itself. Whenever the individual is given practice on some task other than the job itself, the problem of transfer of training is raised.

The Nature of Transfer. The transfer effect may either be positive or negative, *i.e.*, the learning of one act may either facilitate or hamper the subsequent learning or execution of another act. With positive transfer, the individual who has had the opportunity to practice on the transfer task is superior in performance on the subsequent task to the individual who has not had such practice. In the case of negative transfer, the individual who has had this previous practice displays inferior performance on the subsequent task.

Several explanations have been offered to account for transfer, the two most widely accepted ones involving the carry-over effects of general principles and the functioning of identical or similar components.[31] In the first, the general principles are thought of as abstractions about which generalizations to more than one situation can be made. For example, a principle of attack learned in one situation may require that the learner always study the problem as a whole before analyzing it into its parts and attacking each part separately. Such a general principle when applied to new problems may facilitate or retard their solutions. If applicable in the new situation there should be a positive transfer effect. If the principle is inapplicable, whether or not it is so recognized by the learner, the effect should be negative or detrimental in character. For example, an individual given problems to solve not only learns to solve each problem but in addition develops general modes of attack that he carries over to further problems. Thus the individual learns to solve or learns to learn.[38] With coding clerks it is found that the time required to learn successive codes gradually diminishes.

According to the notion of identical components the learning situation and the subsequent job situation have parts in common. The term parts, as used here, may refer to common stimulus characteristics, response components, attitudinal elements, interest tendencies, work habits or techniques, etc. Sometimes even similarity of principles is included, and no distinction is then made between the two explanations of transfer effects. As an example of identical components, a positive transfer should be expected between learning arithmetic and performing the operations of bookkeeping because of the similarity of the arithmetic components.

Negative Transfer. One form of negative transfer frequently encountered is called habit interference. It is exemplified by the difficulty encountered in learning to typewrite by the touch system after having typed for several years with the "hunt and peck" system. This example illustrates the fact that negative transfer is most likely to occur in learning to make a new response to a stimulus or situation to which a different response previously was made.[31] Consider a process operation in a chemical plant where, in response to information given by a dial, the worker adjusts the flow of material by operating a certain mechanical control. Suppose that in training he is given practice on old equipment where the control is a lever whereas on the actual job the control is a wheel. Under these circumstances his performance in utilizing the wheel would be detrimentally affected. On the other hand, if the controls in both the training and job situations were the same but the dials were different, his performance on the job would be less likely to be detrimentally affected.

Evaluation of Gains from Transfer Tasks. Ordinarily the beneficial effects of transfer are measured in terms of how much time is saved in learning a subsequent task as a result of having previously received training on a transfer task. Suppose that as a result of having received training on task A, trainees require only one week to reach a given standard of performance on task B, whereas those not trained on task A require two weeks. It would then be said that experience on task A transferred to task B, which is a desirable consequence. However, suppose the training on task A itself requires two weeks. Then utilizing task A as a transfer task would increase training time from two to three weeks. Since training costs, salaries, time involved, etc., would be increased by 50 per cent it would be concluded that the utilization of task A reduces the effectiveness of the training program even though it does result in positive transfer.

Suppose that performance of the main task involves costs not required by the transfer task. A good example would be pilot training in the air as contrasted with training on a Link trainer. An airplane uses gasoline, requires a higher professional level of instruction, and involves more costly

maintenance. In this case, even though total learning time were increased, there might be a considerable over-all saving not only in dollars but also in lives. Therefore, in order to evaluate the effects of utilizing a transfer task it is not sufficient to take into account just learning time or even level of proficiency attained; gains and losses in other respects must be considered.

Another factor involved in the evaluation of gains from transfer tasks revolves around the time when the gains may be expected to be manifest. Ordinarily it is presumed that if there are any transfer effects they will show up immediately. It appears, however, that under certain circumstances the beneficial effects of transfer do not occur in the initial performance of the main task.[20] It may be that the skills, attitudes, etc., learned in the transfer task are not important in the early phases of learning the main task but contribute primarily in the later phases.

Special problems in transfer arise in connection with synthetic training devices. These are devices which attempt to provide opportunity for practice, and perhaps for guidance, in apparatus which simulate the actual job. The Link trainer, for training airplane pilots, is an example of a synthetic training device. Training devices have been developed for many different jobs ranging all the way from operating a file to aerial gunnery. The cost of manufacture and maintenance of such devices sometimes is fairly high and must be taken into consideration when computing gains.

Evidence concerning the transfer effectiveness of synthetic trainers is ambiguous.[17] In some instances gains are achieved while in others there is no transfer at all. For example, a study of the effectiveness of an elaborate device designed to provide training in gunnery on the ground for fighter pilots proved of little or no value.[13] Pilot trainees received 4 to 25 sessions on the device known as the *gunairstructor* and the amount of such training was not found to be significantly related to accuracy of firing in the air. In this case, even though superficially the two tasks seemed similar, they obviously involved different abilities and different circumstances since the correlation between performance on the gunairstructor and in the air was only .11.

Some Factors in Transfer. The many studies that have been done on the effects of transfer of training clearly demonstrate that the extent of transfer from one task to another is a function of the similarity of the tasks.[31] The more similar the two tasks the greater will the learning of one be facilitated by the prior learning of the other. Unfortunately the degree of similarity of tasks is difficult to determine. Because two tasks involve the same types of materials or activities does not ensure that they are psychologically similar. At present there is insufficient information to justify precise statements as to what constitutes similar and dissimilar

tasks. This means that conclusions concerning the transfer effects of the learning of one task upon another should be made with caution. Special studies to demonstrate transfer empirically are indicated.

There is one important exception to the general rule that the transfer task should be as similar as possible to the main task. This exception arises out of those situations where in the learning of the main task certain factors are at work which slow down learning. Emotion-provoking conditions would be an example of such factors. Jobs which are performed under stressful or dangerous circumstances may be slowly learned. If the learning could be accomplished under conditions where the stress and danger are removed, at least in the initial phases, the rate of improvement should be greater. For example, the job of welding requires that the worker wear an unfamiliar face protector, and the activity of welding is accompanied by flashes, sparks, and noises that may frighten a trainee and hence retard learning. It would be expected that if acquisition of welding skills were accomplished utilizing a simulated welding device without these distractions and disturbances, learning would proceed at a faster rate. After an appropriate level of skill had been achieved, the learner then could be given the actual welding machine to operate.

When transfer-of-training methods are employed, the usual procedure is to have the transfer task easier than the main task. Thus the student pilot is trained first on less complicated aircraft. There is some evidence with perceptual-motor tasks, however, that there is greater transfer from a difficult to an easier version of a task than from an easier to a more difficult one.[22] Such results, however, would be expected only if the difficult task were not so difficult as to present a stressful situation or to result in discouragement from lack of achievement.

Occasionally the situation exists where performance on the main task does not begin until some time after learning on the transfer task is accomplished. This situation may occur as a result of circumstances in connection with the scheduling of various phases of the training program. A bus-driver trainee during his initial classroom instruction may be taught the operations involved in making change and collecting fares using a simulated cash box. Next he is taken out on the road and taught to drive the bus. During the period used in training him in the operation of the bus he receives no practice, either direct or indirect, in making change and collecting fares. Once the driving skill is established he performs the entire job of operating the bus and dealing with passengers. The question arises as to whether transfer effects will be lost during the period when the activity under consideration receives no practice. The results of investigations of this problem indicate that there is perhaps not very much more loss with passage of time than can be accounted for on the basis of simple forgetting.[8]

Transfer from Training to Job. In setting up a training program it is important to give some concern to the relationship between the learning situation and the job situation. One general suggestion frequently made is that the trainee should do his learning under conditions identical with those with which he will be faced on the job. This generalization needs further examination. The specific purposes of training are not identical with those of job performance. In addition to being concerned with production, training is concerned with instilling correct understanding, correct modes of response, correct attitudes, and correct motivation. There may be certain conditions characteristic of the job that would have a detrimental effect upon one or more of these other objectives.

Suppose the performance is one of operating a punch press, and on the job the task must be executed in a noisy, distracting environment produced by the operation of many other machines. Teaching correct procedure would be detrimentally affected under these distracting conditions; therefore training should be conducted in quiet surroundings. Then, after correct procedures have been established, practice should be given under the noisy conditions. A distinction must be made between job factors that facilitate training and job factors that retard training. The argument is not being advanced that training under exact job conditions would not show a high degree of positive transfer, but rather that the greatest positive transfer effect would result from an intelligent discrimination of the effects of job factors upon training, and the use of these factors at different appropriate stages during the learning.

The learning of correct methods of work during formal training is important owing to the possibility of negative transfer. One of the common findings in industrial training is the ease with which incorrect work methods may be acquired by the new worker. Once learned, there may be considerable difficulty encountered in rejecting these methods and replacing them with correct methods. Through negative transfer effects, wrong methods of work may unduly prolong the initial adjustment problems of the new worker when he is placed on the job. A good example of this type of transfer is the classical study of bricklayers by Gilbreth.[23] The method of work learned in formal training had to be unlearned because it differed so greatly from the way the tasks actually were performed on the job.

Transfer from the training period to the actual performance on the job must be conceived as taking place in all aspects of a given job task. There may be either positive or negative transfer effects in methods, procedures, and techniques of work, in the speed of responding, in the accuracy of the performance, in the interest, drive, and urge to respond, in the attitude toward all phases of the work, and in any other characteristic attributable to the job performance. An example of negative

transfer effects involving the rate-of-performing-the-job reactions is that of Gilbreth on bricklayers already cited. He found that trainees instructed under slowed-down conditions learned a set of movements that handicapped them in performing under the normal working speed. It was necessary for them to learn a different set of reactions under the faster rate used in actual bricklaying. It also was found that these earlier reactions retarded the speed of learning of the correct responses.

It must not be thought that these transfer effects are found only between training and job performance of novices; they apply equally well to the training of older workers for new jobs. The problem with older workers is of even greater difficulty because of the potential transfer effects of their old skills, attitudes, and abilities to any new type of work. Established skills, similar in nature to the ones to be learned, may not necessarily show positive transfer to the learning of the new task. Negative transfer effects may result when the more rapid pace characteristics of the older skill are carried over to the new with resultant mistakes and breakdowns. The worker may find it difficult to slow his pace to a rate commensurate with his unfamiliarity with the new tasks. Even when no formal training is given to the older worker before he is transferred to a new job, there must be a transition or breaking-in period during which any interference from his old skills can be overcome and adequate adjustment made.

TRAINING METHODS

Spacing the Learning Effort. One question that arises in connection with training is whether it is better to lump all the practice periods together or to separate them in time. This is the problem of massed versus distributed practice.[31] Endless tasks are monotonous and result in boredom. Breaking the learning into units and distributing the units in time enable the learner to set up a sequence of goals. As each goal is reached, the resulting knowledge and satisfaction from achievement serve to maintain and increase the motivation.

Usually there is a tendency to condense the training into a short time period so that the worker can enter the job as soon as possible. In industry the usual argument offered against distributed practice is that while the trainee is not practicing he is not learning but is still being paid. This point of view ignores the fact that any job is complex and requires learning in more than the specific response patterns required by the job. The periods between practice of the actual job responses is available for instruction in background information, knowledge of the design and maintenance of any equipment operated, safety procedures, proper emphasis upon speed and quality of work, company policies regarding the work, and other similar items which not only contribute to

improving job performance but are basic to establishing a correct attitude toward the job.

While in some cases massed practice gives better results, in general if a given amount of practice is to be allowed, more effective learning occurs if the practice periods are distributed.[31] In one study, for example, fighter pilots were allowed 2,000 rounds in practice firing at a towed target.[13] Some pilots used up their ammunition in four aerial missions, while others took as many as eight. The results, shown in Table 13-3,

Table 13-3. The Effects of Distribution of Practice (Number of Missions) on Accuracy of Aerial Gunnery of Fighter Pilots at the End of Training

No. of missions taken to shoot 2,000 rounds	*% hits on target on last mission*
8	5.3
7	3.8
6	3.0
5	1.9
4	1.0

clearly indicate that the greater the distribution of practice the greater the accuracy of performance at the end of training. In fact, the performance of those who had eight missions was more than five times better than that of those who had only four missions.

It is important to point out that distributed practice not only results in more rapid learning, it also results in superior retention of the activities learned.[35]

Learning the Components of a Task as an Organized Whole. In connection with the learning of a complex skill the question arises as to whether the worker should attempt to master the whole task at one time or whether he should master one element or component at a time and then later organize the several components together into a whole. For example, in learning to make an assembly the learner could go from the beginning to the end of the whole process of making a complete assembly, then follow this with another complete assembly, etc. Or he could learn each of the separate manipulatory responses, practicing each by itself, and then after all of the separate responses were learned he could put them together into the complex assembling operation.

Whether the whole method or the part method is superior depends upon several factors, primary among which is the closeness of the functional relationships of the part activities. In learning a skill the performance of a segment of a complex task by itself is different from its performance in combination with the rest of the responses involved in the

task.[3] For example, sitting behind the wheel of an automobile and rotating the wheel as a small child sometimes does is not the same response pattern as manipulating the steering wheel while driving the car.

The superiority of the whole method comes from the relationships that are naturally found, or which the worker is able to form, among the components.[12] In driving a motor vehicle the judgments about distances of objects, and the differentiation of dangerous objects from others, bear directly upon the operation of the brakes, the throttle, and the steering wheel. The latter activities are a direct consequence of the former, and the perceptions, judgments, and manipulatory responses are all bound together into a unitary reaction. In learning a closely knit series of responses by the part method of learning the coherence between components is not established. The advantage of the whole method increases as the functional integration of the component activities increases in importance.

Some industrial tasks are not made up of closely integrated responses and therefore the part method has an advantage.[5] In carpentry, for example, the various activities are somewhat psychologically independent. The acts of sawing, fitting together materials, nailing, etc., are done separately and therefore each within itself is a functional unit. Since carpentry consists of a number of independent activities, the part method would be utilized in learning the trade.

The part method may have some advantage when certain parts of an act are much more difficult than other parts.[5] Thus in the beginning of learning when the response-patterns are novel and therefore difficult, the part method may prove more effective than the whole method.[10]

Training for Accuracy versus Speed. In many industrial tasks not only is speed of performance important but the avoidance of errors, too, is highly desirable. In typing, for example, good performance not only involves speed but also accuracy. In training workers for such tasks a question arises concerning the relative emphasis that should be given these two factors. A generally accepted principle of training holds that in the initial stages of learning, if emphasis is placed upon accuracy rather than upon speed, ultimate benefits accrue to both. By and large the evidence pertinent to this problem supports this notion.[32]

There are certain cases in which the lack of emphasis on speed may have adverse effects.[23] In some tasks the movements made when performing them at a slow rate are quite different from those made when performing them rapidly. If speed is so de-emphasized that the learner makes very slow motions, the result may be the development of inferior habits of work. As a consequence it is important to introduce some emphasis on speed as early as possible during the learning period. This would appear

to be especially true with tasks that involve rhythmic movements or some patterning in time of the movements.

KINDS OF INSTRUCTIONAL PROCEDURES

Instructional procedures must be as varied as the company and worker's training needs demand. Space is not available for describing all of the types of procedures that have proved valuable. Instead, attention will be directed to a few general forms that currently have a widespread application.

Job-performance Training Procedures. In job-performance instruction the worker is trained while he is performing the tasks required by the job. This category, however, is broader than on-the-job training and apprenticeship training, and includes the instruction given in vestibule schools and company schools where the trainee learns by performing responses similar to those required by the job although not learning in the real job situation.

On-the-job training tends to stimulate high motivation in the learner. In starting immediately on the job the worker attains one of his strongest immediate goals, which is to start producing. In working with the actual tools, materials, and methods of the job he gets the feel of the "real" thing. He recognizes himself as performing the same tasks as other workers around him and gets added security from identifying himself with them. All of these features are positive factors in the personal adjustment of the trainee. It is possible for learning to be detrimentally affected, however, if the worker's introduction to the job is too abrupt and the initial tasks made too difficult for him to comprehend and master readily.

The vestibule school provides facilities for training the new worker before he actually enters the job itself.[15] It has several characteristics which are conducive to high motivation. The worker is going to school on company time. He is working on the same types of machines that he will operate on the job. He gets the opportunity to try out for more than one job if there are openings. His adjustment to the job, following formal training, is made easier. In the school he can automatize to some degree the reactions he must later use, and this proficiency engenders confidence which enables him to adjust better to the unfamiliar and distracting conditions of the actual working environment. A further advantage of the vestibule school is that it is organized around a learning situation rather than around a production situation, and makes possible the practice of many psychological principles of learning which cannot be applied in the actual job situation.

Most of the advantages of the vestibule school also are found in com-

pany schools. In these schools training is usually alternated with practice on the job. After the trainee has learned certain phases of job performance in school he is transferred to the actual working environment where he can use the responses he has learned. When he has demonstrated an acceptable level of proficiency he is returned to the school for further formal training.

In apprenticeship systems the trainee gives his services to an employer at a moderately low rate of pay while the employer is teaching him the trade.[33] Modern systems are highly organized. They contain explicit statements of the particular skills to be developed, units of instruction with promotion from one unit to the next higher one as the apprentice qualifies, ample provision for job experience in all of the skills, and a written contract guaranteeing both the employer and the apprentice a fair return for their contributions.

The Conference or Group-discussion Method. As the name implies, this method requires the participation of the members of a group in a free exchange of opinions, ideas, and criticisms in the solution of a problem.[4,26] Usually the group is small, and hence the method has the advantage of stimulating active participation of all the members. This makes for better mastery and retention of the material learned.

The agenda are usually sufficiently flexible so that participants can suggest problems that are troubling them and that are in need of further thought and analysis. Systematic thinking can be applied to any problems that are presented. As a result of the free interplay of ideas it is usually possible to iron out differences and approach common agreement. This is important where matters of company policy are concerned.

The method is conducive to a high level of cooperation and morale among the participants. Each member has the opportunity of contributing to the program and therefore identifies himself with it. Each sees the necessity of compromise, and each comes away from the discussion with not only a greater understanding of the problems analyzed but also a greater appreciation of the problems and points of view of the other members. These features of the conference method are psychologically basic factors in the development of mutual trust and cooperation.

The method has certain disadvantages. Conferences often fail because there are no definite problems to discuss, or because the problems presented are not of significance to the members. Furthermore, the method is limited to problems that can be dealt with on an abstract, verbal level, and is not an effective procedure for developing specific skills and knowledges. Its most serious limitation is that, as a training device, it is slow and therefore somewhat costly.

The Interview-counseling Method. This method involves a face-to-face interpersonal relationship between the worker and a training officer who

is expertly qualified to assist in the solving of problems of human behavior. By means of an interview, supplemented by other diagnostic procedures such as psychological tests, medical examinations, and biographical histories, the worker and counselor learn the facts about the worker's problem. Then through discussion and counseling the instructor assists the worker to formulate a solution. He further aids the worker in any way which will better enable him to effect the solution such as suggesting stages through which the solution might evolve, providing appropriate training aids and instructional materials, being available for further discussions and counsel, etc. The method is sometimes used for disseminating information and for coordinating company policies, but these are functions better served by the lecture and conference methods, respectively. The primary value of the interview-counseling method is in the solving of specific problems of the individual worker.

The effective use of the method can be illustrated with the case of a bus driver who had experienced a long series of accidents. His performance in other respects was excellent. Knowing that he would be released if he continued having accidents, he sought out the counselor to discuss his problem. With the counselor's help, he analyzed his accidents and discovered that most of them had occurred on the right side of the bus. In addition there were a number of complaints of his closing the door on passengers and of failing to stop for people waiting on the curb. It was suggested that his field of vision be tested. The tests showed that he had a restricted field. Obviously, his field of vision could not be increased, but the possibility of developing new habits of looking to the right side seemed worth exploring. A training procedure was set up whereby the driver was placed in the center of a large semicircle formed by six lights. Whenever a light turned on he had to depress an appropriate key. This required him to turn his head much more than he was accustomed. After his performance in this situation reached a satisfactory level he was assigned to a bus on an easy run and was observed by an instructor. The driver carried over his newly formed habit to the driving situation and his performance was acceptable. He was then placed on his old run and his safety record was found to be adequate. By intelligent application of the counseling method a readjustment was effected which resulted in the worker continuing satisfactorily on his job.

The Lecture Method. Under this heading are included all meetings called among workers for the purpose of presenting information orally. The lecturer is presumed to occupy a position of authority or prestige that entitles him to address the group. The method is frequently used because of the ease and rapidity with which the information can be given to a large group of workers.

The method has several limitations.[18] It is restricted to the type of

training involving the understanding and learning of verbal materials. It is not very successful for disseminating information to workers in jobs at the lower levels of skill because these workers are unaccustomed to concentrating and learning in this type of situation. The success of the method depends too largely upon the ability of the lecturer. It also depends upon authority for getting over the points to be learned, and use of authority sometimes stimulates resistance rather than cooperation on the part of the listeners.

REFERENCES

1. Anastasi, A.: Practice and variability, *Psychol. Monograph*, **40,** No. 5, 1934.
2. Bartlett, F. C.: "Remembering: A Study in Experimental and Social Psychology," Cambridge University Press, 1932.
3. Beeby, C. E.: An experimental investigation into the simultaneous constituents of an act of skill, *Brit. J. Psychol.*, **20,** 336–353, 1930.
4. Bergen, H. B.: Executive training, *Soc. Advance. Mgmt. J.*, **2,** 66–70, 1937.
5. Blackburn, J. M.: The acquisition of skill: an analysis of learning curves, *Ind. Health Research Bd.*, No. 73, 1936.
6. Blain, I. J.: Principles of industrial training, *Occupational Psychol.*, **18,** 1–7, 1944.
7. Blankenship, A. B., and H. R. Taylor: Prediction of vocational success in three machine operations, *J. Appl. Psychol.*, **22,** 518–526, 1938.
8. Bunch, M. E.: The amount of transfer in rational learning as a function of time, *J. Comp. Psychol.*, **22,** 325–337, 1936.
9. Carr, H. A.: Teaching and learning, *J. Genet. Psychol.*, **37,** 189–218, 1930.
10. Coch, L., and J. R. P. French: Overcoming resistance to change, *Human Relations*, **1,** 512–532, 1948.
11. Cox, J. W.: "Manual Skill," Cambridge University Press, 1934.
12. Crafts, L. W.: Whole and part methods with unrelated reactions, *Am. J. Psychol.*, **42,** 591–601, 1930.
13. Crawford, M. P., *et al.*: "Psychological Research on Operational Training in the Continental Air Forces," Army Air Forces Aviation Psychology Research Reports, No. 16, 1947.
14. Dashiell, J. F.: Experimental Studies of the Influence of Social Situations on the Behavior of Individual Human Adults, in Murchison, C. (ed.): "Handbook of Social Psychology," Clark University Press, 1933.
15. Donald, W. J.: "Handbook of Business Administration," McGraw-Hill, 1931.
16. Elwell, J. L., and G. C. Grindley: The effect of knowledge of results on learning and performance, *Brit. J. Psychol.*, **29,** 39–53, 1938.
17. Fitzpatrick, R.: "The Development of a Research Program on Advanced Synthetic Electronic Type Flight Simulators," American Institute for Research, 1950.
18. Ford, A. A.: "Scientific Approach to Labor Problems," McGraw-Hill, 1931.
19. Gagné, R. M., and H. Foster: Transfer of training from practice on components in a motor skill, *J. Exp. Psychol.*, **39,** 47–68, 1949.
20. Gagné, R. M., and H. Foster: Transfer to a motor skill from practice on a pictured representation, *J. Exp. Psychol.*, **39,** 342–354, 1949.

21. Ghiselli, E. E.: Changes in neuro-muscular tension accompanying the performance of a learning task involving constant choice time, *J. Exp. Psychol.*, **19**, 91–98, 1936.
22. Gibbs, C. B.: Transfer of training and skill assumptions in tracking tasks, *Quart. J. Exp. Psychol.*, **3**, 99–110, 1951.
23. Gilbreth, F. B.: "Motion Study," Van Nostrand, 1911.
24. Greene, J. H.: "Organized Training in Business," Harper, 1928.
25. Haire, M.: Some problems of industrial training, *J. Social Issues*, **4**, No. 3, 42–48, 1948.
26. Humke, H. L.: Types of foremen training, *Society Advance. Mgmt. J.*, **2**, 143–148, 1937.
27. Katona, G.: "Organizing and Memorizing: Studies in the Psychology of Learning and Teaching," Columbia University Press, 1940.
28. Kreuger, W. C. F.: The effect of overlearning on retention, *J. Exp. Psychol.*, **12**, 71–78, 1929.
29. Langley, R. W., and J. R. Edwards: Training and selection of workers, *Society Advance. Mgmt. J.*, **1**, 31–36, 1936.
30. Maier, N. R. F.: An aspect of human reasoning, *Brit. J. Psychol.*, **24**, 144–145, 1933.
31. McGeoch, J. A., and A. L. Irion: "The Psychology of Human Learning," Longmans, 1952.
32. Myers, C. S.: Speed versus accuracy in the development of industrial skill, *J. Personnel Research*, **4**, 20–22, 1925.
33. National Industrial Conference Board: Studies in Personnel, Report No. 18, 1940.
34. Postman, L., and R. F. Jarrett: An experimental analysis of "learning without awareness," *Am. J. Psychol.*, **65**, 244–255, 1952.
35. Reynolds, B., and I. McD. Bilodeau: Acquisition and retention of three psychomotor tests as a function of distribution of practice during acquisition, *J. Exp. Psychol.*, **44**, 19–26, 1952.
36. Seashore, S. E., in Gray, J. S. (ed.): "Psychology in Human Affairs," McGraw-Hill, 1946.
37. Spitzer, H. F.: Studies in retention, *J. Educ. Psychol.*, **30**, 641–645, 1939.
38. Ward, L. B.: Reminiscence and rote learning, *Psychol. Monograph*, **49**, No. 220, 1937.
39. Wickens, D. D., G. R. Stone, and R. W. Highland: A study of the retention of electronics fundamentals during basic radar mechanics training, Air Training Command, Human Resources Research Center, *Research Bull.*, No. 52–26, 1952.

CHAPTER 14

Motivation of Workers

Of all the areas of industrial psychology perhaps the most important, and certainly the least understood, is that of motivation of workers. To ferret out the determinants of even some relatively simple behavior of an individual is by no means easy. Indeed, in many cases it is seemingly impossible to understand the bases of a worker's actions. An employee who apparently has been contented on his job for years will suddenly flare up because his supervisor never has greeted him in the morning with more than a perfunctory hello. Well-paid workers will strike for even more wages. A vice president in charge of operations will become disgruntled when the company assigns him a two-door sedan whereas the vice president in charge of sales is assigned a four-door sedan.

GENERAL PROBLEMS IN MOTIVATION

Sources of Difficulty in Understanding Motivation. The lack of understanding of motivation is the result of difficulties in definition and the sheer complexity of the phenomenon. The term motivation is used to refer to so many different aspects of behavior that its meaning is far from clear. Man, in common with other living organisms, has needs and desires that impel him toward certain goals or ends.[33] His behavior is therefore purposive in the sense that it is directed, and hence is organized in reference to particular ends.[36] Motivation might be thought of as the process by which a need or desire is aroused, and a motive thought of as a particular need or desire.

Some writers, however, find it more satisfactory to think of motives as psychic forces within the individual. Presumably these forces initiate and direct behavior. By some these forces are considered to be wholly innate in the individual and by others to be both innate and acquired. Motivation has also been conceptualized in a variety of other ways.[22] Since concepts with respect to motivation are not well formulated, it is impossible at the present time to provide a rigorous definition of the problem. The

best that can be said is that motivation concerns the dynamic processes producing goal-oriented behavior.

The second factor that makes motivation difficult to understand is its sheer complexity. A man quits his job stating that the pay is insufficient. It might therefore be said that his motivation is economic gain. However, on further investigation it turns out that he wants more money so that he can refurnish his home. It seems, then, that his motive is to achieve more gracious family living. But his wife has insisted that home standards need to be raised. Consequently his motivation would seem to arise out of a desire to please his wife. He married this particular girl because she looked like his mother whom he loves. Shall it therefore be concluded that the man quit his job because he loves his mother? Obviously the chain could be continued further and unendingly. Such an analysis, though plausible, is fruitless. The original problem is lost sight of in the very complexity of the situation. Rather, as Krech and Crutchfield have suggested, the best that can be done is to seek the determinants of behavior in terms of the external environment of the individual, his internal physiological states, and his past experiences.[22] Even seeking an adequate and complete description of these three conditions is difficult and complex. Hence it will be necessary to be satisfied with considerably less than a complete description of the motivational determinants of behavior.

The Importance of Motivation in Industry. It has only been in the last ten or fifteen years that industry has given any significant recognition to workers as human beings. While it would be stretching the truth too far to say that in the past the worker has been regarded as a mere machine, it is nonetheless true that behavior was thought of as being determined in a simple mechanical fashion. A given set of circumstances was expected invariably to produce the same kind of reaction. If intensity of illumination were increased 10 per cent, production was expected to show a comparable increase. Sometimes, however, production was found to drop while on other occasions it was found to rise in a striking fashion. This "perverse" character of human behavior was given some cognizance, but it was regarded more in the nature of "error" rather than being lawfully determined.

In more recent times, the "human" aspects of the worker have been subject to much closer scrutiny. As a consequence there has been increased understanding of the dynamic factors underlying behavior, and an increased appreciation of the importance of feelings, attitudes, and motivation. Whereas, in the past, problems in motivation were regarded as being merely problems in discovering adequate incentives—and primarily financial incentives at that—today in industry there is great concern with the multi-motivational determinants of behavior. The worker

is no longer regarded as a compartmentalized creature—the man at home being completely unrelated to the man at work. He is seen as a whole person whose feelings, attitudes, and motivations interact in complex ways with his capacities and with the physical and social environments in which he finds himself.

As a result of this recognition considerable progress has been made toward better understanding of the factors that determine human behavior, and better ways have been developed for dealing with the individual so that he can become not only more productive but also better adjusted occupationally. Business, industrial, and governmental organizations are showing great concern for what is called "human relations." Programs are being developed directly pointed at discovering and capitalizing upon the motivations of workers. The result has been beneficial not only for management but also for the individual worker.

Overemphasis of the Importance of Motivation. Even a cursory glance at current writings in the field of worker problems will reveal a consum-

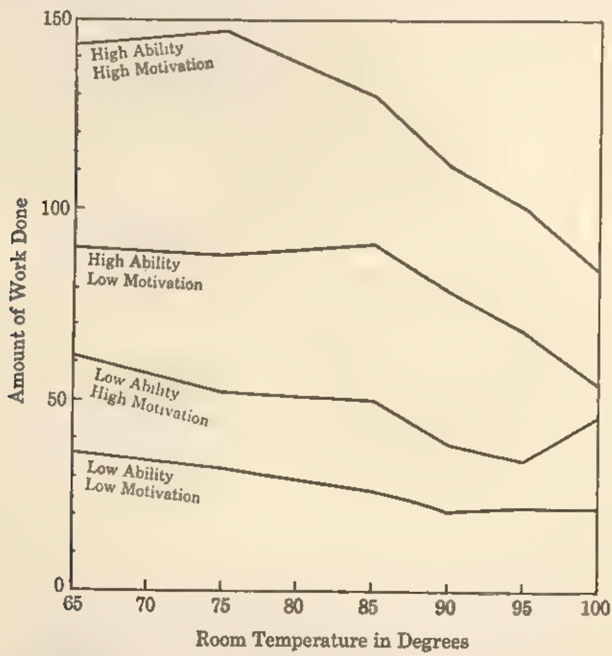


Fig. 14-1. The effects of level of motivation, temperature, and ability upon heavy muscular work.

ing interest in motivation and allied matters. This stress undoubtedly is far too great. The implication is that such factors as ability, learning, equipment design, and conditions of work are of little or no importance.

Essentially the point of view taken is that high motivation will compensate for all manner of unfavorable conditions whether they are within the individual, such as low ability, or environmental, such as noise.

This position obviously is too extreme since the significance of other factors has been amply demonstrated. This is not to minimize the importance of motivation but rather to point out that motivation is only one of many factors that determine workers' behavior. The role played by motivation should be given its appropriate emphasis, no more, no less.

The relative effects of motivation, conditions of work, and ability can be seen in an investigation conducted by Mackworth.[24] This investigator had men perform heavy muscular work under various temperature conditions. In some cases the men worked under low conditions of motivation and in others under high conditions involving rewards, praise, competition, and individually set goals. All men underwent all conditions of temperature and motivation and were divided into the upper half and lower half in terms of their ability to perform the work. The results, shown in Fig. 14-1, clearly show that performance under high motivation is superior to that under low motivation. It is also apparent from these results that high motivation did not compensate for adverse atmospheric conditions. With higher temperatures, performance decreased regardless of whether motivation was high or low. Furthermore, at no point did high motivation overcome the effects of low ability. The point being made is not that motivation is unimportant but rather that other factors may be equally important or even more important than motivation in determining workers' performance.

METHODS FOR STUDYING MOTIVATION OF WORKERS

In most instances it is impossible to study the motivation of workers with the same type of controlled experimental investigation that can be utilized with such factors as methods and conditions of work. Since the motivational determinants of behavior are so complexly interrelated it is difficult to keep all but one factor constant when studying its effects. There are three types of methods for studying motivation: those which infer motivation from behavior, those which involve direct reports from the individual concerning his motivations, and those which utilize the so-called "projective" techniques.[22] None of these methods is entirely satfactory, but all, in one way or another, serve to furnish some information concerning motivation of workers. For each method there are many types of specific techniques and modifications.

Inferences about Motivation from Behavior. In this method inferences concerning an individual's motivation are drawn from the characteristics of his behavior. Thus the individual who takes great care in his work

and who stands admiringly at his final product may be said to be motivated by workmanship. The characteristics of the behavior that lead to inferences about motivation are many and certainly are not clear-cut. They include such characteristics as being goal-directed, continuous, typical, and involving satisfaction or dissatisfaction.[22]

This method, of course, may result in erroneous conclusions. For example, if a worker quits his job it would seem a fair inference to say that he is not satisfied with it. But it might be that his wife wants him to take employment closer to home, or perhaps he has found another job that is more satisfying. Because of the possibility of error, any conclusions evolved through this method should be checked by some other method.

Knowledge about workers' behavior may be obtained in a variety of ways. Kornhauser has summarized these into three approaches: [20]

1. Statistical analyses of labor disputes, grievances, absenteeism, turnover, etc.
2. Firsthand observations of the behavior of individual workers, supervisors' reports and descriptions of workers' performance, etc.
3. Historical analyses of growth and changes in labor unions, analyses of writings and similar expressions of opinion about workers' problems, etc.

Reports from the Individual Concerning His Motivation. Superficially it would appear that the best way to ascertain an individual's motivations would be to ask him. Indeed it is true that a great deal of valuable information can be obtained by this method. The individual can report on his thoughts, feelings, and goals so that in many cases a well-integrated picture of the important determinants of his behavior will result. However, as Kornhauser has pointed out, this method has three important difficulties.[20] First there is the question concerning the individual's willingness to report. It would require rare and most favorable circumstances to elicit statements from an individual that are derogatory to him or are contrary to local mores. Can it be imagined that a rugged logger would complain that the boss hurt his feelings? Secondly, there is the question of whether the individual is capable of reporting his motivations. As will be pointed out later much motivation is unconscious, and hence the individual is not aware of all the forces that really drive him. The third difficulty revolves around the fact that reports are likely to be distorted by recent events. Thus if workers are questioned just after pay day, when the financial pinch is off, they may give little stress to financial motives.

In spite of these difficulties the method of individual report is widely used today as a basis for gaining knowledge about motivational factors. Reports are obtained either by means of an interview or by a printed questionnaire, the latter frequently being termed a morale or job-satisfaction questionnaire. Various forms of interviews and questionnaires

have been utilized.[20] The general advantages and disadvantages of each were discussed earlier in connection with job analysis. The individual may be questioned directly concerning his motives, aims, and goals, or he may be asked about matters he considers of importance or interest, or about his complaints or fears, or his reasons for wanting a particular job, or his reasons for leaving one job for another.

While the interview and printed questionnaire methods may seek information concerning the same factors, they do not necessarily give the same results. This means that to some extent the findings from studies involving individual reports are a function of the method. This is another reason for exercising caution in drawing conclusions concerning motivation when only one method of study has been used. Wedell and Smith, for example, found that results obtained from interviews tend to indicate more favorable attitudes than results obtained from questionnaires.[40] Differences between methods, however, are by no means constant but vary from one question to another.

There is some question concerning the extent to which the results of a study involving individual reports can be accepted. In addition to the three difficulties mentioned earlier, the problem of reliability and validity arises. If, when questioned concerning motivational factors on two different but temporally close occasions, workers give quite different answers to the same questions, then it would be said that the reliability of the questions is too low to be useful. Statistics concerning reliability are by no means extensive, and in general the results indicate the reliability to be moderate but certainly not high.[14]

The problem of validity is more complex but it would be expected that an individual's ratings of motivational factors would be related to some extent with his behavior on the job. Studies of such relationships have given quite discrepant results. Berenberg had 890 factory workers fill out a questionnaire which gave an indication for each person of his general morale, his acceptance of the organization, and his opinion concerning supervision.[1] None of these characteristics was found to correlate at all with absences, tardiness, visits to the medical unit, or merit ratings. Giese and Ruter, on the other hand, did find relationships between job performance and morale scores as determined from a questionnaire.[10] These investigators ranked 25 departments in an organization according to the average of the morale scores of the employees and found that those departments ranking high had somewhat better production, turnover, tardiness, and absentee records.

Projective Techniques. Since direct questioning of an individual may prove inadequate because he does not recognize his true motives, indirect approaches have been suggested.[22] In direct questioning the situa-

tion is structured or organized for the individual and he responds within this set framework. To a direct question such as "What is the most important thing to you in your job?" the individual must respond with words and phrases such as pay, good supervision, etc. If the individual is not aware of his real motives, he is likely, as Haire and Grunes have pointed out, to answer a question such as the foregoing with clichés, stereotypes, or other coherent irrelevancies.[11]

An indirect approach would involve the use of ambiguous or unstructured stimuli such as pictures, meaningless figures, or incomplete sentences. For example, a picture of a worker half facing his machine may be presented and the question asked, "What is this man thinking about?" Since there is no set basis for response, the individual will be forced to project his own personality, thoughts, and motives into the situation in order to respond. Hence the name projective techniques has been applied to such devices.[9] There are no limitations placed upon the response and it might therefore be in terms of some job-connected matter such as "He's going to ask the boss for more money," or some home-connected matter such as "He's worried about his wife's health." The claim is made that projective techniques reveal a great deal about motivation of which the individual himself is not aware.

Table 14-1. Reasons Given by Retail-grocery-store Workers for Liking Their Jobs as Revealed by Direct Questioning and by a Projective Technique

% of reasons		Reason given
Direct questioning	Projective technique	
19	19	Interesting job
17	15	Likes associates
12	3	Contact with customers
12	10	Working conditions
9	17	Wages
7	12	Fits well with habits of life and work
6	2	Easy work
5	2	Fair company
5	5	Job security
3	7	Supervision
2	5	Future advancement
2	2	Autonomy
1	1	Union protection

Projective techniques are relatively new devices and have not been widely used in the study of workers' motivations. Comparative studies indicate that they give somewhat different results than direct questioning. Haire and Gottsdanker interviewed retail-grocery-store workers asking direct questions and in addition obtained responses by projective techniques.[12] The results shown in Table 14-1 indicate that certain factors received different emphasis under the two methods.

SOME CHARACTERISTICS OF MOTIVATION

Motivation Is Complex. The motivational determinants of behavior are not related in any simple summative manner, but rather the relationships are very complex.[23] The interplay among a common group of factors is such that in different situations different factors play the dominant role. Under one set of circumstances a given condition might result in changes of one kind, whereas under other circumstances the results might be quite different. The interactions among motives, conditions of work, and various aspects of the social milieu are far more important in the determination of human behavior in the industrial situation than any single motivating, environmental, or social condition. The difficulties encountered when attempting to explain or predict the behavior of workers may be illustrated by the following two examples.

A group of employees was engaged in work involving considerable eyestrain that experimental tests had indicated could be reduced by blue-green lighting.[26] When this type of light was introduced in the actual working situation the output of the men workers increased, and complaints of fatigue were reduced. With the women, on the other hand, output fell off, and complaints, dissatisfaction, and absences increased. When the cause of this unexpected reaction on the part of the women was investigated, it was found that they felt that the new type of lighting made them look "simply ghastly." The introduction of an apparently favorable condition of work therefore may not necessarily achieve the desired ends if it brings opposing or antagonistic motives into play.

Fairchild studied men engaged in various metal trades in four factories.[7] The work in factory A required the least amount of skill, that in factory B more, that in factory C still more, and that in factory D the most. Interviews with the workers in these factories indicated that in factory A the major source of satisfaction to the workers was the wages they earned. The conditions of work and the job itself contributed very little. In factory B the conditions of work supplemented by wages were the chief sources of satisfaction. The job itself and conditions of work were about equally important sources of satisfaction at factory C. At factory D the job itself was the outstanding source. These findings indicate that where

skill is outstanding it is an important source of satisfaction to the workers and tends to depress the importance of wages and conditions of work. On the other hand, when the skill required is of a low order it becomes unimportant, and wages and conditions of work increase in importance as sources of satisfaction. Thus, although skill, conditions of work, and wages may each operate to produce satisfaction in the workman, when operating in conjunction their effects are not simply added together but rather are functionally dependent upon the interrelationships among the factors.

Some Motivation Is Not Recognized by the Individual Himself. One characteristic of motivation that has already been mentioned is that the individual is by no means always aware of the true nature of that which impels him to action.[22] He may be cognizant of his motives or he may be completely unaware of them. This characteristic of motivation is frequently overlooked by management and supervisors. It does not appear reasonable that an individual does what he does without knowing why. Yet this is true in many instances.

An example can be drawn from studies conducted at the Western Electric Company.[32] A worker counseling program had been instituted in the company wherein each employee was regularly interviewed by a professional counselor who sought to uncover the sources of worker dissatisfaction and to correct the maladjustment. One girl complained to her counselor about her foreman and was able to document her case in considerable detail with regard to specific unfavorable actions on his part. Later in the same interview the conversation turned to the girl's home life. Her situation there was most unfortunate because of the rigid discipline of an overbearing stepfather. He restricted her activity far more than was warranted and made unreasonable demands of her. In describing her stepfather, the girl said that physically he looked very much like her foreman. The picture now fell into place both for the counselor and for the girl. The girl had unconsciously transferred to her foreman the unfavorable characteristics of her stepfather.

But how can the girl's specific and detailed complaints about her foreman be explained? This can be understood when it is recognized that the actions of another person are not either all bad or all good. In many instances the behavior will be innocuous or ambiguous and hence there will be considerable margin for interpretation. Because of motivation that she did not recognize, the girl interpreted the ambiguous actions of her foreman as unfavorable. The implication to be drawn from the facts of unconscious motivation is that an individual's statements concerning his motivation should not be accepted at face value.

Motivations Change. Another characteristic is that the motivations of each individual change from time to time even though he may con-

tinue to behave in the same way. This characteristic is easily recognized in the larger frame of reference of the life span of the individual. Thus neither the young child nor the aged person is expected to direct a large portion of his attentions to members of the opposite sex. Similarly, to some extent there is an appreciation of the relatively quick changes in motivation as manifest in the fluctuations of moods. It is not regarded as strange that a man will be motivated to avoid social contacts at breakfast time, while in the evening he will actively seek such contacts. However, changes in motivation between these two extremes tend to be ignored.

Consider the following example. A man seeks a job in order to provide the financial means for supporting his family. In inquiring about the possibilities of employment, he asks about wages and security of employment. He goes to work and maintains a high level of production in order to strengthen his job security. As time goes by he finds himself steadily employed and so is reasonably assured in regards to future security. With this off his mind he becomes more deeply engrossed in his work and finds that the job is interesting in and of itself. Motivation directed toward workmanship now serves to maintain his proficiency at a high level. With the passage of the years the job becomes more and more an integral part of his personal and social life, and high job proficiency then becomes just as necessary to him as eating and sleeping.

The individual's motivations change as he grows and develops not only physically but also intellectually, emotionally, and occupationally. Furthermore, the industrial situation in which the employee works is seldom static for long periods of time. New methods of work are introduced, supervisors change, and indeed, the entire physical environment of his work place may change. Add to these the changes in social and political views of the nation together with variations in economic conditions, and wonder can be expressed at whatever stability in motivation does exist.

Individuals Differ in Their Motivations. As might be expected, individuals differ in their motivations. Two people may do the same thing for different reasons or different things for the same reason. One student goes to college in order to fulfill his parents' wishes while another attends because he desires to become a lawyer. One man attempts to achieve promotion by increasing the amount of work he does and another by increasing quality. In part, motivation grows out of the individual's experiences. Since different individuals have different experiences it is not unexpected that their motivations are different.

The wide variation that exists among workers can be illustrated by a study conducted by Wyatt and Langdon.[43] They had 325 women factory

workers rank 10 motivational factors according to importance. The percentages of workers assigning the various ranks to the item of opportunity for promotion are shown in Table 14-2. It will be noted that the

Table 14-2. The Percentages of Workers Assigning Various Ranks to the Item "Opportunity for Promotion" When Comparing It in Importance with Nine Other Items

Rank	1	2	3	4	5	6	7	8	9	10
% of workers	6.8	8.7	9.6	16.4	13.9	12.4	13.0	11.5	5.6	2.2

item appears at every rank, indicating that promotion is far more important for some workers than for others. Furthermore, there is no one rank in which the item receives a distinctively large proportion so that it is impossible to assign to it a single value that will indicate its importance as compared with the other nine items.

The important implication of individual differences in motivation is that the same incentives and appeals will not operate equally well with all workers. Suppose that two workers ignore the same safety regulation that requires the wearing of safety goggles. In correcting the workers the best procedure for one might be to threaten him with separation because he is motivated for financial gain in order to support his family. For the other, who has no such obligations, the appeal might be to job pride since he is young and wants status among his fellows.

Motivations Are Varied and Specific. Not too long ago most managements operated on the thesis that the only motivation of workers was a desire for financial gain. Incentives and other means developed for motivating workers were based upon this premise. However, it is clear that there are many different types of motivational factors that determine the performance of the worker on his job. While financial gain is important to most workers it is by no means the sole factor. Strikes called on the basis of a wage dispute may be settled on other than financial grounds. Promise of an increase in remuneration to an already highly paid executive might operate not as a financial incentive but rather as an indication of recognition for outstanding performance.

In subsequent pages the results of several studies of worker motivation will be described. It will be seen that there are a whole variety of factors which are important in addition to financial gain and even economic security. The viewpoint that there are only one or two basic drives that determine behavior is untenable. Men do sacrifice their lives

that others may live, and they do voluntarily leave higher paying jobs for lower paying ones. The goals to which an individual aspires are many, and so are his motivations.

FACTORS IN THE DEVELOPMENT OF WORKERS' MOTIVATIONS

Factors in the Working Situation. The attitudes and feelings that workers develop on their jobs enter into the determination of their motivations. If management appears fair and working conditions good then one pattern of motivations will emerge. If pay schedules are good but quality of supervision is poor, then another pattern will result. In any event, since motivation is to some degree a function of the individual's experiences, those experiences which are job-connected may be expected to play an important part.

Perhaps the most striking demonstration of the influence of attitudes and feelings on job performance comes from the famous "Hawthorne experiment" carried out at the Hawthorne plant of the Western Electric Company.[32] The investigation grew out of the failure to find the expected drop in output of a group of workers when they were subjected to decreased intensity of illumination. The change in illumination apparently brought about a change in the attitudes of the workers so that the deficient working conditions were compensated for by an increase in effort. From these findings the investigators realized that only by means of a prolonged research program could results be obtained that would clarify the causal conditions. Consequently six girls who assembled relays were placed in a segregated experimental room, called the test room, for a period of over two years and were subjected to a variety of standardized conditions. The conditions of work to which they were exposed are recorded in Table 14-3. As shown in the table, from the beginning to the end of the experimental period the girls showed a steady increase in production.* Even under conditions when output might have been expected to drop, increases were manifest. During the period of the investigation there was an increase in contentment among the girls, a greater interest in the job and in the working situation, and a reduction in absences of about 80 per cent.

On the basis of repeated interviews and observation of the girls the investigators were led to conclude that the increase in output was due to a change in attitude. Before any changes in the working situation were made the girls were consulted, their opinions as to the effects of the changes were solicited, and they worked under conditions free from hard-and-fast supervision. It appeared that the workers developed a new

* Estimated from charts given by Whitehead.[41]

Table 14-3. Working Conditions and Relative Productivity of Girls Engaged in Relay Assembly in the Hawthorne Experiment

Experimental period	Duration, weeks	Working days per week	Daily working hours	Rest periods	Scheduled rest pauses, total min.	Special features	Relative output per week
1	2	5½	8¾	0	0	Regular department	100
2	5	5½	8¾	0	0	Test room	101
3	8	5½	8¾	0	0	Group rate pay	105
4	5	5½	8¾	2	10	Two 5-min. rests	109
5	4	5½	8¾	2	20	Two 10-min. rests	112
6	4	5½	8¾	6	30	Six 5-min. rests	113
7	11	5½	8¾	2	25	Rests 15 min., A.M.; 10 min., P.M.	116
8	7	5½	8¼	2	25	4:30 stop	123
9	4	5½	7¾	2	25	4:00 stop	125
10	12	5½	8¾	2	25	Check period	124
11	9	5	8¾	2	25	Saturday morning off	123
12	12	5½	8¾	0	0	No rests	122
13	31	5½	8¾	2	25	Rests	131
14	9	5	8¾	2	25	Saturday morning off	133
15	31	5½	8¾	2	25	Rests	135
16	4	5½	8¾	2	25	Operators changed positions	136
17	25	5	8	2	25	4:15 stop, Saturday morning off	138
18	15	4½	8	2	25	4:15 stop, Friday afternoon and Saturday morning off	140 140
19	15	4½	8	2	25	Operators returned to original positions	138
20	25	5	8	2	25	4:15 stop, Saturday morning off	138
21	—	4	8	2	25	4:15 stop, Monday and Saturday morning off	138

social orientation that was the major determinant of their behavior in the industrial situation, a new attitude that was not adversely affected by the experimental changes introduced.

Factors outside of the Working Situation. If the worker's attitudes and feelings growing out of the working situation affect his behavior on the job, it is not unreasonable to suppose that attitudes and feelings developed in his home life and other off-the-job activities would also have important effects. Hersey intensively studied a small group of workers engaged in different occupations.[15] When they were in a positive and happy mood, their output was increased by some 2 per cent, and when they were disturbed it was reduced by 7 per cent. In a study of employees in a British cocoa works, reported by Hall and Locke, the home life was found to have very important effects.[13] The workers did not divide their life into two watertight compartments, one inside the factory and the other outside of it. The two were closely bound together so that the troubles and joys of home life could not be put aside when reporting for work in the morning, nor could factory matters be dropped when returning home after work. An unsatisfactory home life was reflected in reduced effectiveness on the job.

FINDINGS FROM STUDIES OF WORKER MOTIVATION

Numerous studies have been made of worker motivation. These studies have employed various methods and have been aimed at different objectives. Their multiplicity makes it difficult to organize the findings into a coherent whole. For purposes of presentation typical studies will be considered under the following categories: stated aims and desires of the worker; workers' complaints, grievances, and fears; problems in which workers express interest; and reasons given for leaving the job.

Stated Aims and Desires of the Worker. Strong determined the aims of workers by noting down their desires as they have been reported in various publications.[34] He lists the following aims:

1. Steady employment: no discharge without cause, seniority
2. Job requirements: clear instructions, definite allocation of responsibility, freedom to do job in own way, being consulted about changes in job, good equipment and materials
3. Working conditions: protection against accidents and illness, proper heating, lighting, ventilation, sanitation
4. Wages: equitable, higher, sufficient to provide for comfort, proper differential according to skill
5. Hours of labor: shorter, vacation
6. Escape from fatigue, strain, exhaustion, monotony
7. Treatment as an individual, respect for opinions, voice in control of welfare conditions, individual freedom, freedom to consult and make suggestions, confidence of superiors
8. Enjoyment of work: knowledge of results and larger affairs of business
9. Voice and free will in determining work conditions, sense of responsibility
10. Grievances satisfactorily adjusted
11. Opportunity to rise on merits
12. Decent boss, real leader, fairness not sympathy
13. Approval of fellows and public, prestige
14. Recreational facilities, leisure
15. Savings, ownership of home
16. Insurance against risks of life as in accidents, sickness, old age, death
17. Abundant life, more education for self and children, best existence and happiness for family

Strong does not pretend that the above is a complete list of workers' motives, but simply offers it to give some notion as to their important desires.

Hersey presented groups of employees with 14 items relating to the working situation and asked them to check the four most important ones and the four least important ones.[16] The results, shown in Table 14-4, indicate that steady employment, fair adjustment of grievances, amount of pay, working conditions, and safety are important factors to the employee, whereas a voice or share in the management, employee stock

subscription, medical and dental care, and methods of pay are relatively unimportant.

Significant differences exist in the stated desires of union and nonunion workers. Whereas a large proportion of union men considered adjustment of grievances as being important, relatively few nonunion

Table 14-4. Important and Unimportant Factors in the Working Situation as Indicated by Workers' Responses on a Questionnaire

Aspect of working situation	Important, %		Unimportant, %	
	Union	Nonunion	Union	Nonunion
Steady employment	65	93	3	10
Fair adjustment of grievances	80	24	6	27
Amount of pay	49	51	0	3
Working conditions	49	45	3	3
Safety	57	21	6	7
Chance for promotion	28	47	13	7
Type of man in charge	18	38	28	34
Insurance systems and pensions	18	36	16	10
Hours of work	13	23	3	21
Voice or share in management	13	6	78	69
Chance to show initiative	5	6	47	34
Employee stock subscription	5	2	100	93
Medical and dental service	0	6	72	17
Methods of pay	0	2	25	65

Table 14-5. Rankings of Importance of 10 Job Factors Made by Workers in Three Occupations

Job factor	Sales	Clerical	Mechanical
Security	4	2.5	1
Advancement	2	2.5	2
Type of work	1	1	3
Company	3	4	4
Coworkers	7	5	5
Pay	5	7	7
Supervisor	6	6	6
Hours	9	9.5	8
Working conditions	8	8	10
Benefits	10	9.5	9

men did so. Differences of this kind between groups is a common finding. Jurgensen, for example, found that variations in motivation were related to marital status, number of dependents, age, education, skill, and occupation.[18] His findings for sales, clerical, and mechanical workers are shown in Table 14-5. While all three groups agree that "benefits" are quite unimportant, type of work is the most important factor for sales and clerical workers, whereas for mechanical workers security is the most important factor.

Workers' Complaints, Grievances, and Fears. By means of the interview method, Centers compiled the specific complaints of workers who were dissatisfied with their jobs.[3] The findings, presented in Table 14-6,

Table 14-6. Major Complaints of Workers Who Were Dissatisfied with Their Jobs

Complaint	White-collar workers	Manual workers
Inadequate remuneration	26	18
Insecurity	9	14
Work too hard	2	18
Ambitions	13	5
Little chance for advancement	11	8
Poor hours of work	6	8
Lack of freedom	6	5
Work monotonous	9	2

show striking differences between white-collar and manual workers. To a considerable extent these differences arise out of the nature of the jobs. Thus manual workers complain about the demands of the job far more often than do white-collar workers. To some extent, however, the differences are probably due to the types of individuals involved. Eckerman analyzed the grievances filed by machine-shop and foundry workers.[6] These findings, given in Table 14-7, are similar, though by no means the same as those reported by Centers for manual workers.

Uhrbrock has suggested that the worker is motivated by three major fears, viz., fear of losing his job, fear of loss of power to earn resulting from sickness or accident, and fear of an impoverished old age.[37] Kornhauser and Sharp found that, on a list of worries presented in a questionnaire, most factory workers checked the item "fear of loss of job," and on a list of dislikes most of them checked the item "possibility of layoff." [21]

Table 14-7. Analysis of 766 Grievances Filed by 377 Machine-shop and Foundry Workers

Type of grievance	*% of grievances*
Pay and wages	30
Job and work	28
Seniority	10
Union business	9
Company business	7
Matters for collective bargaining	6
Promotion and transfer	5
Discharge and reinstatement	3
Vacation	2
Total	100

The importance of these fears in factory workers has also been confirmed through interviews with workers by Hall and Locke.[13] In addition they found that fear of novelty, ridicule, and disapproval are considered important by workers. The study of employees' fears would seem to be a fruitful source of information concerning the determinants of workers' behavior. Unfortunately most of the interest appears to have been centered around fears based on economic considerations. Hall and Locke's study indicates that an investigation of fears arising out of the social situation would be fruitful.

Problems in Which the Worker Expresses Interest. In a plant of the Western Electric Company a system was introduced whereby the employees were interviewed relative to their work and their personal problems.[30] Cooperation was ensured since the employees were aware that their statements would be held confidential. The comments, both favorable and unfavorable, were classified according to the topic to which they referred. The following topics of interest to the employee were found:

Absence	Lighting	Social contacts
Advancement	Lockers	Steady work
Aisles	Material	Temperature
Club for employees	Monotony	Thrift
Dirt	Noise	Tools and machines
Fatigue	Payment	Transportation
Floor	Placement	Vacation
Furniture and fixtures	Restaurant	Ventilation
Hospital	Safety and health	Washroom
Hours	Sanitation	Welfare
Interest	Smoke and fumes	Working space

It is important to note the wide variety of topics of such interest to the worker that he discussed them in the interview. It also is interesting to note that very few of the interests have to do with economic security.

Richards classified the topics mentioned by 243 factory workers in five interviews of the kind described above.[31] The proportion of workers mentioning each topic and the proportion of their remarks that involved grievances are shown in Table 14-8. The findings are similar to those

Table 14-8. The Percentage of Factory Workers Making Remarks about Various Topics, and the Percentage of These Remarks That Involve Grievances

Topic	% of workers mentioning	% of remarks that are grievances
Safety	37	57
Speed of work	17	93
Bonus	4	100
Wages:		
Rate	23	75
Distribution	6	100
Hours of work	18	91
Steadiness of work	8	68
Transfers and promotion	10	75
Supervision:		
Fairness and impartiality	18	34
Efficiency	11	62
General nature	42	45
Environment:		
Associates	12	83
Women in plant	6	86
Light, heat, ventilation	21	82
Cleanliness	9	95
Personnel and welfare activities	10	50
General plant procedure	25	82
Materials	7	100
Equipment	14	91
Attitude toward work	16	58

observed in the study just cited. The coefficient of correlation between the proportion of workers mentioning the topics and the proportion of remarks that were grievances is $-.52$. This means that the less important a topic is in terms of the number of workers mentioning it, the more likely

will the remarks about it be unfavorable. This cannot be due to fear on the part of the workers to mention grievances, since 76 per cent of all remarks were grievances.

Reasons the Worker Gives for Leaving the Job. When an employee voluntarily leaves a job it means that he is not entirely satisfied with it. By means of an exit interview the reasons given by employees for leaving a job of their own accord can be systematically obtained and recorded. This type of information often discloses important determinants of behavior that otherwise would go unnoticed. The results of an investigation of aircraft-factory workers during World War II can be cited as typical.[29] Half of the reasons for leaving were job-connected and half were personal. The specific job-connected reasons in order of frequency were placement, desire for another job, general dissatisfaction, wages, shift, work load, supervision, and working conditions. The personal reasons in order of importance were health, child care, transportation, and home responsibility.

FACTORS IN JOB SATISFACTION

Estimates of the incidence of satisfied workers in the general employed population are in fairly close agreement.[8,17,35] It appears that a majority of workers, about two-thirds, are in general satisfied with their jobs. Although statistics such as these are useful in giving an over-all picture, they obscure certain important characteristics. Job satisfaction has many different points of reference, and few workers indeed are satisfied with all aspects of their jobs.[13] The over-all measurement of job satisfaction may obscure differences in degree of satisfaction in specific areas, compensatory adjustments that occur with variations in working conditions, and other related factors. The satisfactions the worker derives from his job are not only many and varied but are also highly sensitive to change. Interviews with workers have indicated that there is considerable variation in satisfaction from day to day and from week to week, and even hourly changes have been discovered.[13]

Job Satisfaction as a Function of Occupational Level. It is ordinarily believed that individuals working in the higher occupational levels are more satisfied with their jobs than those in the lower levels. More persons in professional and managerial jobs are thought to be satisfied with their work than those in office jobs, and so on down to unskilled jobs. Investigations of the relationship between occupational levels and extent of satisfaction do not indicate this to be wholly true. Super found that there is an increase in the proportion of satisfied workers from the lowest manual jobs up through the skilled occupations.[35] At this point a break occurs, so that, with the lowest office occupations, less satisfaction is

again in evidence. From this level upward, increases in occupational level result in further increases in extent of satisfaction. This suggests that workers make a distinction in their minds between manual and non-manual jobs, and that at the higher levels of each type more satisfaction is manifest than at the lower levels.

As might be expected, changes in occupational level—promotion and demotion—also affect job satisfaction. Super found that 67 per cent of those who had not changed level were satisfied, but only 41 per cent of those who were reduced in level were satisfied. Related to the problem of promotion and demotion is that of the relationship between job satisfaction, on the one hand, and the difference between the worker's aspiration and achievement on the other. Super also found that, of those individuals whose jobs were exactly of the nature to which they aspired, 76 per cent were satisfied, whereas, of those individuals whose jobs did not measure up to what they aspired, only 34 per cent were satisfied.

Caste as a Determinant of Job Satisfaction. Not only are there variations in satisfaction with broad occupational groupings, but even within jobs that are roughly similar the worker makes a differentiation as to desirability. Although in part the ideas of differences between jobs are based on differences in wage payments, an equally powerful factor is the standing among his fellow workers that the job gives to the man, *i.e.*, his caste. Williams, who worked among men in factories, mines, and railroads in an attempt to get a firsthand picture of workers' motives, has reported that there is a vast social gap between the individual who holds a "swell" job and one who holds a "bum" one.[42] For example, a laborer transferred to the millwright gang in one plant received an increase in pay of only 2 cents per hour, but the shift immediately changed his social status; he was congratulated by his fellow workers, and his wife and family were "received" where before they had not been.

The causes of caste are many and often obscure. Cleanliness of work is important, particularly among women workers.[13] If a man's work is connected with that of a laboratory technician, he may bask in the reflected glory of his superior. Even among employees carrying on exactly the same work at the same pay there may be observable differences in social status. Seniority, no doubt, plays a part. The position of the machine or desk at which the work is done and its relative newness or oldness may be important.

Within a given group, workmanship may set some individuals above others.[32] The worker who displays exceptional skill is looked up to by his fellows. As Fairchild noted in interviews with men in metal trades, workers derive considerable satisfaction from their skill.[7] Even workers engaged in such simple occupations as making boxes and tending ma-

chines may enjoy the work for its own sake.[13] The de-skilling of jobs may result in dissatisfaction since the individual feels that his capacities are not being adequately utilized.[39]

It seems likely that this pride in work grows out of the worker's identification of himself with his job. As Williams puts it, workers have "a sense of ownership in a piece of property called their jobs." [42] This personal identification with the job is manifest in a variety of ways. Railroad men refer to their engines as the "old girl" with no small amount of warmth and friendliness. Hall and Locke found that factory workers do the same sort of thing, christening their machines with such proper names as "Joey" and "Maggie." [13] Apparently, when given the opportunity, workers develop close associations with their jobs that not only result in increased satisfaction but also in good workmanship. That such a spirit should arise is not surprising when it is remembered that the worker spends almost as much time on his job as he does in active life at home.

Age and Job Satisfaction. Age has been reported as related to job satisfaction and also as not related to it. Results obtained by Super indicate that the lack of agreement is due to the fact that satisfaction does not vary directly with age, but rather shows cyclical changes with age.[35] His findings indicate that individuals aged 25 to 34 and 45 to 54 appear less satisfied than those at other ages.

Job Satisfaction as a Function of Financial Incentives. The fact that methods of paying employees for their work vary so widely makes their evaluation difficult. Much has been written concerning the good and bad aspects of different payment procedures, but for the most part the arguments are based on rationalization rather than on systematic investigation. For instance, a bonus plan might be said to be inadequate because good workers become discouraged when they see poor workers receiving an equal amount of money. Much research remains to be done in this field. A few typical results will be considered below.

A survey by the National Industrial Conference Board indicates that profit-sharing plans are seldom put into force on the basis of employees' suggestions.[28] When such plans operate successfully they appear to increase output, improve morale, and decrease labor turnover. In many instances, however, they do not work out satisfactorily. In analyzing the reasons why profit-sharing plans were given up by various organizations, the Conference Board found that in about one-third of the cases there was unfavorable reaction on the part of the employees, and in another third, dissatisfaction on the part of the company. The remaining third gave no explanation or reported a variety of reasons. Workers' dissatisfaction with profit-sharing plans appears to stem from an understandable dislike of an irregular income.

Although payment on a piece-work basis brings about an increase in output and wages and a reduction in costs, it also results in a voluntary restriction in output.[38] Rightly or wrongly, most workers appear to feel that if their earnings are too high the rates will be cut so that they will have to work harder and harder to earn the same amount of money. Through interviews with workers, Mathewson found that this voluntary restriction of output is far greater than is generally realized.[25] In some situations the motivation to conform to a restricted level of output is so great that workers may deliberately not report work accomplished if it brings their production over the accepted level.[32] Besides restricting output in the belief that he is protecting his rate, the employee also does it to postpone layoff and to steady employment. It is brought about not only by the worker himself but also by pressure from other workers and sometimes even under orders from superiors. Grievances against management, resentment against speed-ups, discouragement, and monotony quite naturally play a part in the restriction of output, but in the main the stabilization of earnings appears to be the important consideration.

Quality of Supervision and Job Satisfaction. The conclusion drawn from the Hawthorne experiment, that the relationships between employees and the management are very important in the productivity and happiness of the worker, has been corroborated by other investigators.[13] As in the Hawthorne study, Kornhauser and Sharp found that working conditions were relatively unimportant as a factor in employees' attitudes.[21] In two departments in the same plant in which the work was identical, only 29 per cent of the employees in the department with the better physical conditions of work had favorable attitudes, whereas in the other department 71 per cent of the workers had favorable attitudes. The unfavorable attitudes brought about by poor supervision spread to unrelated things. It is apparent therefore that a superficial examination of workers' complaints may give an erroneous picture.

McGregor and Knickerbocker have suggested that workers' objections to changes in the working situation are not so much due to the changes themselves as to the method of instigating them.[26] Since these changes affect the workers they should be understood by them as being necessary and reasonable. This is just what occurred in the Hawthorne study. The workers never felt that a change would be made that would work against their welfare. As a consequence, even presumably unfavorable conditions did not reduce production since the workers understood the experimental nature of the program and its ultimate objectives.

The evidence clearly indicates that worker satisfaction is established through improving employer-employee relationships so that the worker will be led to believe that his problems will be given due consideration

and that he is an important part of the organization. Obviously such morale can only be developed if the worker has some feeling of security in his job.

FRUSTRATION

Since motivation involves goal-directed behavior it will be well to consider the circumstances that lead to failure to achieve a goal, and the types of behavior that the frustrated individual manifests. Much of the worker's behavior which is not understood becomes clear when it is interpreted as the behavior of a frustrated person. For example, the behavior of a worker who becomes less and less concerned with damage to company property becomes meaningful if it is known that he is disgruntled with his employer because he failed to receive the increase in pay he felt he deserved. A consideration of the causes and consequences of frustration will throw further light upon the problem of motivation.[5]

Sources of Frustration. There are three types of situations which may prevent an individual from achieving his goal. First of all there may be a conflict between motives. The choices in modern society are seldom simple. If the piece-rate worker abides by the social mores of the group, restricts his output, and thus is accepted by his fellows as a "good guy," then he cannot earn the money he would like to refurnish his home. He cannot be both a "good guy" and a high earner. If the salesman accepts a promotion to a managerial position which carries higher pay and status, then he will have to give up the many friendly and satisfying contacts he has with his clientele and perform administrative activities which he finds boring. The satisfaction of one desire may mean the thwarting of another. In some instances there will be more than two goals and thus the frustration may be intensified.

A second situation leading to frustration is some type of personal limitation on the part of the individual. The limitation may be either physical or psychological. A 275-lb. man cannot obtain the coveted job of bus driver because he cannot fit into the space provided for the driver. The line supervisor promoted to a higher-level administrative position becomes frustrated because he does not have the intellectual capacity to cope successfully with the new problems. Psychological limitations are not necessarily intellectual but may involve deficiencies in personality characteristics. An individual may not be temperamentally suited to carry the responsibility that goes along with the supervisory position to which he aspires.

A third situation that causes frustration is some environmental block. The block may be either in the physical or in the social environment. Thus the skilled worker is frustrated in his attempts to do the quality of work he desires because of inadequate tools and equipment. Similarly

the mechanic in dirty overalls will be uncomfortable in the company cafeteria largely populated by the office staff. He will be thwarted in his desire to have a hot lunch convenient to his place of work.

Real and Imagined Sources of Frustration. The source of frustration may be either real or imagined. The conditions that thwart purposive behavior may actually exist, or they may be phantasies of the imagination. One salesman may be frustrated because there is little market for his product, while another may be frustrated because he falsely believes no market exists and hence does not exert the needed effort. Whether the cause is real or imagined the frustration is a real one for the individual concerned. Psychologically it makes little difference whether there really is no market or whether the salesman just believes there is none. In either case he is frustrated in his desire for high sales.

It is too easy to accuse the frustrated individual of indulging in unrealities, pointing out that the sources of frustration are false. To make this interpretation indicates a lack of insight into the motivational dynamics involved and prevents or retards effective steps from being taken toward the readjustment of the individual. Furthermore, many situations are anomalous. The different interpretations that may be applied to the same state of affairs may be equally sound. For example, workers may complain that there is little opportunity for promotion in an organization. Management may respond by saying that this is not true since each year 5 per cent of line workers do receive promotions. Who is to say that such a promotion rate is small or large? It will depend upon the individual and his system of values. As a consequence it is necessary to be cautious in categorizing a source of frustration as real or imagined in a situation that allows honest differences of opinion.

Consequences of Frustration. The consequences of frustration are many.[27] The individual who fails to achieve his goal may react in a variety of different ways. The kind and intensity of the reaction will depend upon the nature of the individual, the strength of the motivation, and the source of the frustration. A worker who has been brought up in a culture that does not condone emotional demonstrations may have strong internal reactions that continue over a period of time, while one who has been brought up in the opposite type of culture may curse and throw his tools about and then manifest no further reaction.

Where there are strong motivations and highly prized goals, reactions to frustration are likely to be strong. However, there is by no means a perfect correlation between intensity of reaction to blocking and the strength of the motivation. Sometimes failure to achieve a goal that is only mildly wanted will result in strong reactions. Finally, the intensity of reaction is related to the source of the frustration. Workers are more likely to accept what is to them an undesirable situation from a super

visor who is considered a "good boss" than from one who is not so considered.

The individual who fails to achieve a highly desired goal may generalize his dissatisfaction to other related areas. Katz and Hyman point out that workers display a "spill-over" of dissatisfaction to areas where there is no intrinsic dissatisfaction.[19] They compared the attitudes of workers in a plant who were resentful over a wage dispute with the attitudes of workers who had no such dissatisfaction. All other plant and community conditions were approximately the same. The first group of workers were not only less satisfied with their jobs and working conditions but also with the total community situation.

The consequences of frustration are not necessarily bad. The individual who fails as a salesman may take a job as a clerk and so rechannel his efforts that he achieves success on the new job. Frustration is likely to lead to an intensification of effort. This increase in effort may not be directed toward the original goal but rather may be expended in other directions where success is more likely to be achieved.

MODES OF REACTION TO FRUSTRATING SITUATIONS

Reactions to frustrating situations, while of many kinds, tend to fall into several types. These various modes of behavior are by no means completely distinct from one another. Furthermore, the frustrated individual is likely to manifest several different types of reaction, in many cases successively manifesting one, then another. The modes of reaction that will be considered here are acceptance of the situation, restriving, withdrawal, disorganization, seeking substitute goals, rationalization and projection, identification, and aggression.

Acceptance of the Situation. An individual who fails to achieve the goal he is seeking may simply give up trying and accept the situation. This type of reaction to frustration is more passive than active. The employee who asks for an increase in salary and whose request is refused may continue to perform his job and live in about the same way as before.

Some psychologists hold that there can never be such a passive reaction to frustration. The argument is advanced that frustration necessarily and inevitably produces some kind of significant change in behavior. This point of view has many adherents, but proof for it is very difficult to obtain. Certainly in many instances the thwarted individual does not react in a way that is observably related either positively or negatively to such thwarting. Therefore it would have to be maintained that reaction to frustration may be quite disconnected with the original frustrating situation either in time or in kind of behavior manifested. For example, it might be argued that the employee who does not receive the

raise he covets may not react until several years after the refusal, and then may react by beating his wife. While in some instances such apparently disconnected events are judged inferentially to be associated with a given frustrating situation, the empirical evidence used to justify the inferences is far from convincing. Consequently it is consistent with the facts to hold that on some occasions frustration does result in passive acceptance of failure to achieve the goal.

Restriving toward the Original Goal. As indicated earlier, one of the consequences of frustration is an increase in intensity of activity. This renewed activity may be directed to other ends or toward the original unachieved goal. Subsequent topics will deal with the former type of reaction, which essentially is of a compensatory sort. It is certainly true that in a large number of cases when an individual fails to achieve some particular goal, he reacts by increasing his efforts and reorganizing his activities so that his chances of success are increased. For example, a worker who seeks upgrading to a higher skilled job but fails to pass the requisite proficiency test for such upgrading may attend night school in an endeavor to build up his knowledge and skills. Similarly the manager of a plant who is dissatisfied with the level of effectiveness with which it operates may call in industrial engineers to help him develop plans for improving the operation.

Restriving toward the original goal is likely to be a healthy and forceful way of solving a frustrating situation. It may involve a reanalysis on the part of the individual with respect to the goal itself and to the means he has utilized in striving to achieve it. Obviously in certain instances restriving with respect to an unachievable goal is fruitless.

Withdrawal. One solution to frustration is simply to withdraw, retreating from the conflict situation. The individual may attempt to forget his difficulties and resign himself to the fact that his goal cannot be achieved. However, it is unlikely in this case that complete forgetting will occur, particularly when the motivation is strong and the goal highly desirable. The frustrating experience may be repressed and then not directly manifest itself in conscious experience. Under such circumstances the problem still remains a problem for the individual, but may be manifested in forms of behavior apparently unrelated either to the goal or to the source of frustration. Thus the worker who fails to obtain a desired raise in pay may complain not about management being niggardly but about hours of work being too long.

For the frustrated person, the real world in which he lives is not wholly satisfactory since it does not permit him to achieve his desired ends. Consequently he may seek retreat in a different and more favorable world—one that he himself constructs. Sometimes he indulges in autistic thinking—thinking on a plane of unreality as in daydreaming. In other

cases he may seek to change the way the world looks to him by imbibing alcoholic beverages. Such attempts to withdraw from the frustrating situation are not likely to result in satisfactory solutions.

When the individual is unable to resolve the frustration and is unable to withdraw completely from the conflict situation, mental changes, sometimes of a more or less permanent nature, may result. One common outcome of such a conflict is boredom. The worker finds his work uninteresting and while he would like to stop must nevertheless continue for fear of losing his job. This conflict leads to pessimistic reverie. Objects, activities, and people that he has hitherto found to be stimulating and interesting become neither exciting nor engaging. He becomes less alert and less responsive to his environment, and he feels that he is less intelligent than formerly.[4] Continued exposure to such circumstances may lead to semipermanent or even permanent changes in the individual's personality.

Disorganization of the Individual. Under the stress of frustration the individual may suffer some disorganization. Under very unfavorable circumstances the emotional breakdown may be of a major sort. More commonly it will be of a minor nature, manifesting itself in intemperate language, throwing down tools, etc. There may even be a regression to infantile modes of reaction such as weeping. Whatever the reaction of this sort may be, it is inappropriate to the situation and at best only provides a means for "letting off steam."

Seeking Substitute Goals. Failing to achieve the ends he desires, the individual may obtain release from the tensions built up through frustration by channeling off his energies in other directions. The worker who does not achieve the success he desires in his job turns to model railroading or gardening in his leisure hours. The individual thus compensates for his failure to achieve one goal by seeking another goal where success is more readily attained. In a sense this is a withdrawal or escape, but in many cases the compensatory activity may be a highly desirable form of adjustment. The individual may achieve real successes with benefits to himself, his family, and his employer. The person who fails as a salesman but becomes a highly proficient office worker is an example.

Some types of substitute goals, however, may lead the individual into activities that are so time-consuming or attention-demanding that he is led away from his main course. Hobbies that in and of themselves are harmless, and even potentially constructive, may distract the person from the main task of effectively performing his job and supporting his family in an adequate fashion.

Rationalization and Projection. One common reaction associated with frustration is the attempt on the part of the individual to find excuses for his failure. The unproductive worker who goes to great lengths in

pointing out how poor his health has been exemplifies this type of behavior. Excuse-making of this kind is termed rationalization. It is not a process of reasoning based on the manipulation of sound facts, but rather a process whereby the individual seeks to absolve himself from the onus of personal inadequacy. An explanation of failure is offered in order to sustain the individual's self-respect. Many rationalizations, of course, are innocuous.

In seeking to exonerate himself, the individual may project to some other person or group of persons the cause of his misfortune. The object of projection may be specific or general. The worker may feel that his supervisor is holding him down, or he may attribute his difficulty to management or even "they." Certain elements of danger are inherent in projection since it provides a mechanism by means of which delusions or false beliefs of persecution can be developed. These delusions become strengthened by continued failure or simply by rumination, so that, finally, the individual or group to whom the failure is attributed becomes in and of itself a source of threat and a basis for further frustrations. The worker who does not receive the promotion he believes he warrants now perceives management as the cause of his failure to achieve other goals. His failure to win his girl's affection is attributed to his bosses who tie him to his desk for endless hours and release him exhausted at the end of the day.

Identification. Since an individual who does not succeed in achieving his goals suffers a blow to his ego, it is not surprising to find that chronically failing people seek to identify themselves with some group that has succeeded or at least promises success. This is particularly true when the group in question is actively attacking or is antagonistic to the sources of the individual's frustrations. Thus 20 years ago the frustrated middle class in Germany flocked to Naziism because it promised relief and indeed retribution. So a worker whose strong desire for job security is not satisfied by such guarantees as management offers may seek membership in an organization which otherwise would not attract him and which promises to protect him from his employer and in fact is antagonistic to employers.

Aggression. Thwarted in his desires, the individual may respond with aggressive behavior. Such behavior is very likely to be directed toward what he believes to be the source of his frustration. Failing to obtain praise for his efforts the worker may sabotage his machine. On some occasions aggression may come out in the form of severe violence. However, since aggressive behavior is not generally condoned, due to social pressure, most aggressive tendencies will be manifest in such minor ways as gruffness, indifference, carelessness, and malicious gossip.

Manifestations of aggression in these subtle and roundabout ways are of particular importance since they are so apt to go unrecognized as symptoms of potentially serious underlying maladjustments.

IMPROVEMENT OF MOTIVATION THROUGH WORKER ADJUSTMENT

From the foregoing discussions it is clear that workers' dissatisfactions stem from many sources, and these dissatisfactions are reflected in adverse ways not only in their job performance but also in their home and social lives. It is obvious that motivation and dissatisfaction are related. With the lessening of workers' dissatisfactions there should occur an improvement in the character of their motivation. It is equally obvious that some sources of dissatisfaction cannot be dealt with by management. The problem may be beyond management's capacity or area of responsibility. Unskilled workers cannot expect to be paid $20 per hour, nor can persons in the construction industry expect to work only indoors. Nevertheless there are many sources of maladjustment and dissatisfaction that are job-connected and with which management can deal.

Approaches to the Problem of Reducing Workers' Dissatisfactions. There are two approaches that can be utilized in attempting to reduce workers' dissatisfactions. One involves discovering the sources of dissatisfaction and attempting through administrative action to remove or change the conditions that give rise to the dissatisfaction. The other approach has quite a different orientation. It involves dealing with the dissatisfaction at the personal level, seeking through a counseling process to resolve the motivational and adjustment problems of the individual worker. It should be obvious that ordinarily both procedures in one form or another are used informally by supervisors. A good supervisor will be alert in detecting working conditions that disturb his subordinates and will seek to remedy these conditions. At the same time he will discuss with his subordinates any individual problems that are bothering them.

In some organizations these approaches are utilized in highly formalized procedures.[32] A special department may be set up to collect information concerning sources of dissatisfaction either by means of interviews with workers or through questionnaires. This information is collected and integrated in some central agency and the results transmitted to top management as a basis for determining needed administrative actions such as changes in safety regulations, pay schedules, hours of work, etc. Or a special department whose members are trained professional counselors may seek to assist workers to satisfactory resolutions of their problems through personal counseling.[2] Both of these approaches, in-

formation collecting and personal counseling, will be considered in the following sections.

The Nature of Information-getting Programs. The purpose of these programs is to obtain directly from workers information concerning the causes of their dissatisfactions. Interviews are conducted with workers or questionnaires are filled out by them personally. It is obvious that precautions are necessary in order to reassure the workers that their complaints will not be held against them. As a consequence interviews are conducted on a highly confidential basis, and questionnaires are ordinarily filled out anonymously.

The data collected are classified according to the expressed complaints. Tabulations of the complaints together with commentaries and comments are then passed to top management. The information is used as a basis for setting policies and procedures designed to improve working conditions, and pertinent data are passed down to appropriate officials for use in supervisory training programs.

Evaluation of Information-getting Programs. Workers generally react quite favorably to programs of this kind. In one evaluation of an interviewing program some 80 per cent of workers were favorably impressed with the procedures.[32] Since in the interviews approximately half of the workers' comments about their jobs and working conditions were unfavorable, it seems likely that the workers were generally unafraid of possible adverse consequences to themselves.

In many instances the remarks of workers do not provide the concrete information concerning unfavorable conditions that is ordinarily desired and expected.[10] Much valuable information is gained concerning workers' points of views and attitudes, but many times the comments are too vague to be of great use. Sometimes complaints are made about conditions, such as sanitation, that actually are wholly adequate. At other times workers mention conditions that are well known to management and can be ascertained without recourse to a formal program of interviewing or questioning workers.

Information-getting Plans in Relation to Supervisory Authority. One important characteristic of information-getting plans is the manner in which ordinary lines of authority are by-passed. Information is obtained from the worker not by his immediate supervisor but rather by personnel operating from an entirely separate jurisdiction. Furthermore, the information does not pass through the usual supervisory hierarchy but is transmitted directly to a section of top management.

The matter of by-passing supervisors has generally received little attention, and the problems that arise from it have not been sufficiently explored. Ordinarily it is dismissed from consideration because the reactions of supervisory personnel to information-getting plans have not

been too unfavorable, and because it is believed that the information obtained by such methods can be collated and distributed to appropriate departments with greater dispatch. Furthermore, such a procedure tends to allay any fears on the part of the workers that their complaints will be held against them by their immediate supervisors, and thus it provides a means for uninhibited discourse.

From a purely objective point of view it would appear that the foregoing reasoning presupposes that supervision and supervisory channels must necessarily be inadequate. More specifically it may be said that the following two assumptions are made. First, owing to the particular social situation existing between supervisors and workers, the latter cannot or will not approach the former with problems that have a highly personal reference. Secondly, information obtained from workers by lower supervisory personnel is inadequate and is not effectively passed up the line to higher administrative levels.

It follows that, if these assumptions are true, the information-getting program is simply a device that is added to the present organizational structure to increase the morale and adjustment of workers, and at best only gets at the problem indirectly. In such a case management must be held at fault for inaccurate selection and training of supervisors. On the other hand, if the assumptions are false then the information-getting program assumes a much less important role. If the supervisor has the appropriate talents and background to command the respect and confidence of the worker there will be good supervisor-worker relations, and many interpersonal conflicts can be resolved immediately and at the level where they arise. There would then be no problems that the worker could not discuss with his supervisor because the worker would have no fear of reprisal. Of course the supervisor would have to be thoroughly trained to evaluate the information and determine what portions of it should be passed up for action at higher levels.

The foregoing criticism is not to be interpreted to mean that an information-getting program has no place whatsoever in an industrial organization. In initial attempts to improve the motivation and adjustment of the workers it will be a most helpful implement in providing an immediate and acceptable procedure for workers' complaints to be heard, and in providing some factual material to be used in supervisory training and in the correction of inadequate plant conditions. But when a sufficiently adequate and sympathetic supervisory system is established, its role as a fact-finding procedure and escape valve for workers' emotions should be reduced to a minimum or become nonexistent.

Even with a good supervisory system in operation continuance of information-getting programs may be necessary. Under the best conditions, human frailty and error will be manifest in any well-developed

supervisory system. Continuance of a modified and somewhat restricted program of interviewing or questioning workers by an independent group furnishes a system of checks and balances needed in any social organization. Indeed, if for no other reason than the fact that it is a function of top management, a system of this kind may well be the means for keeping alive in the minds of top management the problems of personnel at lower levels in the organization. Regardless of its advantages the information-getting program demands a split in supervisory authority and responsibility, and unless the division is clear-cut and willingly accepted by supervisory personnel, the plan will not be wholly effective.

The Nature of Personal-counseling Programs. The purpose of personal-counseling programs is to seek resolutions to workers' problems not through changing working conditions, etc., but rather through reorientation of the worker by changing his attitudes and goals. Whereas information-getting programs provide the worker some relief from psychic tensions, they necessarily deal with more superficial problems and problems that are largely job-connected. It was pointed out earlier that interviewing and questionnaire may not get at true motivational factors. Indeed, from interviewing programs it is apparent that the manifest content of workers' complaints—that which they actually complain about—may not actually be the source of their dissatisfaction.[32] Therefore it is necessary to get at the latent or underlying factors by other approaches.

While information-getting programs are restricted to job-connected factors, the evidence is clear that many maladjustments among workers arise from home and social relations outside the plant. Problems of this sort can be dealt with by personal-counseling programs. It is beyond the scope of the present discussion to consider the details and techniques of psychotherapy, though such matters obviously must be understood if a complete analysis of personal counseling is to be gained. However, two matters do need consideration, the manner in which the difficulty is resolved, and the responsibility for the resolution.

In the counseling situation in industry, it is most appropriate for the counselor to utilize a nondirective technique in his interviews with workers. He provides the stimulus for the initiation of the conversation and leads it only in the sense that he tries to get the counselee himself to bring out pertinent and important matters. In the resolution of a difficulty the first step is to provide a clear description of the problem. The counselor never attempts to state the problem directly, but rather, by indirect methods he seeks to have the counselee do so. In many instances the solution is immediately apparent to the counselee once the problem is clear to him, and the tension is thereby relieved. In other instances, where the solution is not so easily reached, the counselor again resorts to in-

direct methods to bring the counselee around to discover and state the most likely solution.

Both in the clarification of the issues and in the development of the solutions, although the counselor may immediately see the problem and the means for its resolution, he does not directly suggest them to the counselee. He uses such knowledge only to direct the thinking of the counselee so that he is able to discover them for himself. Solutions arrived at in this manner appear to be more adequate and more satisfying to the individual than those given to him directly through the authority of the counselor. They are more likely to be accepted, and therefore result in more consequential action. In addition the counselor can never wholly understand all of the factors in the situation. The individual's values and those of his family and friends must be considered in connection with solutions to many problems. The individual's needs and goals, too, are important. Therefore the responsibility for the solution must be placed on the individual himself and not upon the counselor.

In addition the question may well be raised as to whether management can justifiably dictate to those it employs the ways in which they should act with reference to matters that are of a highly personal nature to them. To do so would be going beyond any conceivable moral obligation of management. What right does management have to suggest to a worker how he should deal with his wife, what his relations should be with his church, and how he should finance the purchase of an automobile? Indeed, ordinarily management wants no part of such problems. However, it is within the management's realm to provide the means by which the worker can arrive at his own solutions. Since with nondirective counseling techniques, the counselor neither states the problem, evaluates the factors involved, nor indicates the solution, it is apparent that he, as a member or representative of management, cannot be said to have interfered with the worker's personal problems.

Evaluation of Personal-counseling Programs. While it is true that information-getting programs require that the comments of the workers be held confidential, a very special relationship exists between the interviewer and the worker when personal problems are being dealt with, as in the counseling program. It is the sort of relationship that exists between the doctor and his patient or the lawyer and his client. The interview records are the confidential property of the counselee and of the counselor and can be revealed in part or in whole only with the consent of both parties, and then only after careful consideration and under certain circumstances. Therefore any information obtained by the counselor from the persons being interviewed can at best be passed on only in very general form, if at all. The counseling is an end in and of itself. An interviewing program, however, in addition to offering some solution

to the worker's problem, yields data that can be used as a basis for administrative action.

The problems dealt with in a counseling program require personnel with professional training and background that is not to be found among supervisors. Supervisors should understand the importance of the psychological adjustment of the workers and of themselves, and should realize that personal counseling is not just for the trouble makers and the neurotics, but rather that all individuals have problems that need resolution, and that the resolution is likely to result in greater satisfaction, more effective work, and higher levels of production.

REFERENCES

1. Berenberg, R. E.: Socio-psychological factors in industrial morale: I. The prediction of specific indicators, *J. Social Psychol.*, **36**, 73–82, 1952.
2. Cantor, M.: "Employee Counseling," McGraw-Hill, 1945.
3. Centers, R.: Motivational aspects of occupational stratification, *J. Social Psychol.*, **28**, 187–217, 1948.
4. Davies, A. H.: The physical and mental effects of monotony in modern industry, *Brit. Med. J.*, **2**, No. 3427, 472–476, 1926.
5. Dollard, J., *et al.*: "Frustration and Aggression," Yale University Press, 1939.
6. Eckerman, A. C.: An analysis of grievances and aggrieved employees in a machine shop and foundry, *J. Appl. Psychol.*, **32**, 255–269, 1948.
7. Fairchild, M.: Skill and specialization, *Personnel J.*, **9**, 28–71, 128–175, 1930.
8. Fortune Survey: *Fortune*, Jan., 10, 1947.
9. Frank, L. K.: Projective methods for the study of personality, *J. Psychol.*, **8**, 389–413, 1939.
10. Giese, W. J., and H. W. Ruter: An objective analysis of morale, *J. Appl. Psychol.*, **33**, 421–427, 1949.
11. Haire, M., and W. F. Grunes: Perceptual defenses: processes protecting an organized perception of another personality, *Human Relations*, **3**, 403–412, 1950.
12. Haire, M., and J. S. Gottsdanker: Factors influencing industrial morale, *Personnel*, May, 2–10, 1951.
13. Hall, P., and H. W. Locke: "Incentives and Contentment," Pitman, 1938.
14. Hardin, E., H. G. Reif, and H. G. Heneman: Stability of job preferences of department store employees, *J. Appl. Psychol.*, **35**, 256–259, 1951.
15. Hersey, R. B.: Rate of production and emotional state, *Personnel J.*, **10**, 355–364, 1932.
16. Hersey, R. B.: Psychology of workers, *Personnel J.*, **14**, 291–296, 1936.
17. Hoppock, R.: "Job Satisfaction," Harper, 1935.
18. Jurgensen, C. E.: Selected factors which influence job preferences, *J. Appl. Psychol.*, **31**, 553–564, 1947.
19. Katz, D., and H. Hyman: Industrial morale and public opinion methods, *Intern. J. Opin. Attitude Research*, **1**, No. 3, 13–30, 1947.
20. Kornhauser, A.: Psychological studies of employee attitudes, *J. Consult. Psychol.*, **8**, 127–143, 1944.

21. Kornhauser, A. W., and A. A. Sharp: Employee attitudes, *Personnel J.*, **10**, 393–404, 1932.
22. Krech, D., and R. S. Crutchfield: "Theory and Problems of Social Psychology," McGraw-Hill, 1948.
23. Lewin, K.: Frontiers in group dynamics, *Human Relations*, **1**, 1–140, 1947.
24. Mackworth, N. H.: "High Incentives versus Hot and Humid Atmospheres in a Physical Effort Task," Medical Research Council of Great Britain, 1947.
25. Mathewson, S. B.: "Restriction of Output among Organized and Unorganized Workers," Viking, 1931.
26. McGregor, D., and I. Knickerbocker: Industrial relations and national defense: a challenge to management, Massachusetts Institute of Technology, Department of Economics and Social Science, **2**, No. 7, 1941.
27. Miller, N. E., *et al.*: Effects of frustration, *Psychol. Rev.*, **48**, 337–366, 1941.
28. National Industrial Conference Board: Profit Sharing, *Studies in personnel policy*, No. 2, 1937.
29. Palmer, D. L., E. R. Purpus, and L. O. Stockford: Why workers quit, *Personnel J.*, **23**, 111–119, 1944.
30. Putnam, M. L.: Improving employee relations, *Personnel J.*, **8**, 314–325, 1930.
31. Richards, J. R.: Interviewing industrial employees, *Personnel J.*, **9**, 281–289, 1930.
32. Roethlisberger, F. J., and W. J. Dickson: "Management and the Worker," Harvard University Press, 1938.
33. Stratton, G. M.: "Man, Creator or Destroyer," G. Allen, 1952.
34. Strong, E. K.: "Psychological Aspects of Business," McGraw-Hill, 1938.
35. Super, D. E.: Occupational level and job satisfaction, *J. Appl. Psychol.*, **23**, 547–564, 1939.
36. Tolman, E. C.: "Purposive Behavior in Animals and Men," Century, 1932.
37. Uhrbrock, R. S.: Attitudes of 4,430 employees, *J. Social Psychol.*, **5**, 365–372, 1934.
38. Viteles, M. S.: "Industrial Psychology," Norton, 1932.
39. Walker, C. R., and R. H. Guest: "The Man on the Assembly Line," Harvard University Press, 1952.
40. Wedell, C., and K. U. Smith: Consistency of interview methods in appraisal of attitudes, *J. Appl. Psychol.*, **35**, 392–396, 1951.
41. Whitehead, T. N.: "The Industrial Worker," Harvard University Press, 1938.
42. Williams, W.: "Mainsprings of Men," Scribner, 1925.
43. Wyatt, S., and J. N. Langdon: Fatigue and boredom in repetitive work, *Ind. Health Research Bd.*, No. 77, 1937.

CHAPTER 15

Social Factors in Industry

Among the various factors that determine a worker's behavior are the influences of other persons. The social environment plays a vital part in regulating his actions, molding his attitudes, and directing his motivations. The individual worker in turn makes his contribution to the social environment. His actions, attitudes, and motivations influence the behavior of others. The members of a group, then, interact in a dynamic fashion with each other. The character of the interactions and the effects they have on the individual are a function of the nature of the group and of the individuals who comprise it.

The social forces that operate upon the worker are both powerful and manifold. While the physical environment is a factor in determining the worker's performance, he is relatively tolerant of it. On the other hand, he is very sensitive to his social environment since his interpersonal relations are intimately related to his desires and goals. In the institution for which he works there are various formal and informal organizations that have a conditioning effect upon his actions. To various extents his behavior is determined by his home, his church, his union, his political party, his community, and the many other groups of which he is a member. In terms of the sheer number of groups to which he is exposed, the worker's social environment contributes significantly to the determination of his behavior.

GROUPS AND ORGANIZATIONS

Definitions of Groups and Organizations. There are various kinds of aggregations of persons, and these aggregations vary in their characteristics and importance as determinants of behavior. Some may have little influence upon behavior. Thus an individual may belong to the class of persons termed employees. His membership in this class, in and of itself, has little effect on his behavior. Rather he is considered to be in this class because he possesses certain characteristics and behaves in certain ways.

Aggregations of this kind are unimportant for the present discussion since they are not primary determinants of behavior.

By comparison the individual may be a member of Department Y of Company Z. This aggregation of individuals will markedly influence his behavior. He will work cooperatively with his fellows, he will quarrel with them, his lunch habits will be influenced by theirs, etc. His superior will give him orders and he in turn may give orders to others. Furthermore, he and the other members of the department recognize that they are different from the persons in Department X. They work in a different place, they have different duties and responsibilities, they have a different supervisor, and they believe that the members of Department X are a bunch of lazy loafers.

A social group, then, differs from a simple aggregation of individuals in two respects.[17] First, the members of a social group are in dynamic interaction with one another, and secondly, they perceive themselves as members of a particular group and are so perceived and reacted to by others. The members of a social group have common needs and goals. They are subject to the same types of pressures and have similar experiences that bind them together and differentiate them from other groups.

One special kind of group is termed an organization.[31] An organization is a social group in which the members are differentiated with respect to their functions in connection with the task of achieving a common goal. They have a mutual responsibility in regard to the achievement of their goal, but different members contribute in different ways. Thus a group may or may not have leaders, but if it does it is an organization since some of its members play different roles than do others. In a highly structured organization, such as a business or industry, the functions of the various members may be clearly delineated. On the other hand, in certain informal organizations, such as a clique among workers, the differentiation of roles might not be clearly defined. It is difficult to state the exact point at which a group becomes an organization. The distinction is in terms of differences in roles played by the workers, and disagreement arises in setting up a criterion for determining how different roles must be to, in fact, be considered different.

Organizations and Morale. One characteristic that is helpful in understanding groups and organizations is morale. The term morale is a much used and much abused word in describing workers' behavior. It refers to the adequacy with which a group functions and to its unity and solidarity with respect to its purposes.[17] Blum has pointed out that it is frequently confused with job satisfaction, and while job satisfaction is related to morale it is not the same phenomenon.[2] An individual may dislike his work and still display high morale since he is at one with his fellows with respect to the common ends they are seeking to achieve

through their coordinate efforts. Conversely, a worker may like his job and yet have low morale since he does not feel identified with his fellows as a group nor is he convinced of the desirability of their purposes.

The orderliness of a group or its productivity is not necessarily a sign of its morale. A team of workers may be highly organized with respect to the assignment of functions among its members and may carry out its duties rapidly and well. At the same time the morale displayed by the team might be described as low. Krech and Crutchfield have listed the signs or criteria of high morale as follows: [17]

A tendency for the group to hold together not merely as a result of external pressures but rather through internal cohesiveness

A lack of tendencies of its members to divide into smaller antagonistic factions

An ability of the group to adapt to changing circumstances and to handle internal conflicts

A feeling of belongingness among the members of the group

A communality of goals among the members of the group

A positive attitude of the members with respect to the objectives of the group and to its leadership

A desire on the part of the members to retain the group and a regard for its positive value

A group of assembly-line workers who deal only with some small part of a large and complicated machine may not see the import of their efforts and hence develop low morale. Similarly the members of a union may have low morale because they have lost faith in their leaders' ability to obtain better working conditions and pay for them. Morale in an industrial organization may be low even when the workers recognize the importance of their jobs, respect their supervisors, etc., but do not feel that management accepts them as belonging to the organization. The inability on the part of the worker to identify himself with his organization as a result of some managements' point of view that workers are not an integral part of the organization but rather "hired hands," undoubtedly accounts for much of the lack in morale. Under such circumstances management would not significantly improve morale, and hence motivation, by improving working conditions, character of supervision, or interest value of the job.

GROUP AND ORGANIZATIONAL STRUCTURE

Groups differ from one another in a variety of ways and a number of systems have been devised for describing and classifying them.[12] They not only differ in their purposes but also in their structure. By group structure is meant the nature and characteristics of the interrelationships among the members and the distribution of functions among them. Two groups may have the same purposes and yet their structure might be

quite different. For example, two labor gangs might be assigned the same task of loading gravel into barrows, transporting it to another place, and dumping it. Suppose one gang consists of "old hands" while the other is largely comprised of new workers. Because of the experience of the first group the foreman would only have to give general directions, and the men themselves would more or less automatically take assignments in shoveling or wheeling. As a result of their skill and knowledge the men would pace their own work and would rest when necessary. In the inexperienced group the foreman would have to make the work assignments, give specific orders, and constantly direct the work.

Formal and Informal Structure. Almost all operating organizations, whether governmental or private, business or industrial, are given a

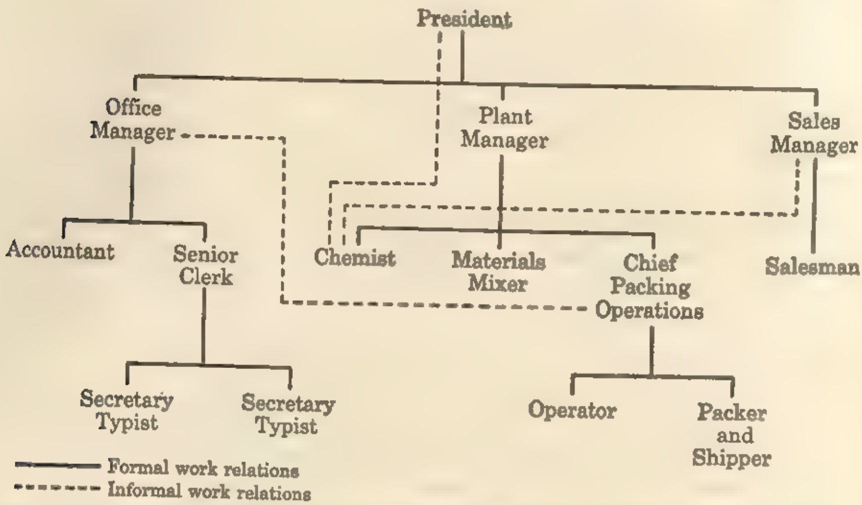


FIG. 15-1. Formal and informal work relations among the members of a small food-manufacturing company.

formal structure. The purpose of such structuring is to prescribe responsibilities and duties so that the functions of the organization can be accomplished in an orderly fashion. Frequently this formal structure is presented in the form of an organization chart which shows the interrelationships among the various jobs that comprise the organization. In such charts the head of the organization, representing the most general level of administration, is placed at the top, and persons of lesser responsibility and having more specific functions are placed below, all individuals connected appropriately by lines to show the flow of authority. An example of a typical organization chart is shown in Fig. 15-1. This chart represents the formal organization in a small food-manufacturing company.

The solid lines indicate the formally prescribed work relations among the individuals comprising the organization

As Stogdill has pointed out, the variables that define a formal organization are responsibility variables—the work the individual is expected to do—and formal interaction variables—the other persons with whom a given individual is expected to work.[31] However, a worker might perform functions somewhat different from those prescribed and deal with persons in the organization other than those formally designated. In the food-manufacturing company mentioned above, some persons wholly on their own performed functions not formally assigned to them and interacted with other persons in the organization outside of set channels. For example, both the president of the organization and the plant manager were interested in product development and worked with the chemist. Similarly, the office manager found it convenient to work directly with the chief of packing operations on matters concerning orders and billing. These ex-officio relations are shown in Fig. 15-1 by means of dotted lines.

An informal structure may then be superimposed upon the formal structure of an organization. This informal structure may be quite different from the formal structure. In point of fact the formal structure may bear little resemblance to the actual characteristics and functioning of the organization. Stogdill gives the variables defining an informal organization as work variables—the tasks an individual actually performs, and informal interaction variables—the other persons with whom a given individual actually works.[31]

Kinds of Factors Determining Group Structure. The structure of a group as indicated by the interrelationships among its members and the functions they perform is determined by many factors. These factors interact in complex ways so that it is difficult to attribute a given structure to one factor rather than to another. Furthermore, the interrelationships among the members of a group and the roles they play in it change from time to time. The structure of the group may be altered as its membership changes and the conditions under which it operates are modified. In the following paragraphs consideration will be given to some of the major determinants of group structure: the degree to which a group perceives itself as a group, and the purposes, size, autonomy, homogeneity, and leadership of the group.

The Degree to Which a Group Perceives Itself as a Group as a Factor Determining Its Structure. Perhaps the most basic factor determining the structure of a group is the extent to which its members perceive themselves as belonging to the group. The line employees of a given organization may each individually consider themselves merely as persons who work for that organization along with a number of others. Such a group would have very little structure, at least of a permanent kind.

Their goals are not common goals but rather individual goals. On the other hand the line employees may see themselves as the most important and the only real productive agents in the organization, whose activities management wishes to hamper, and whose interest management is against. As a result of its beliefs such a group might well become highly integrated, develop new goals, and set up a formal organization, all designed to protect it from real or imagined dangers.

The Purposes of a Group as a Factor Determining Its Structure. It is obvious that the nature of a group's task, its purpose, is an important factor that determines its structure. The activities of some groups require a close integration of the activities of all of its members. The long-distance flight of an airliner requires the coordinate and cooperative efforts of the pilot, the copilot, the navigator, and the flight engineer. While each aircrew member has his own particular functions to perform, these functions are not performed in isolation. The course that the navigator plans will depend upon information given him by both the flight engineer and the pilot. The navigator in turn advises the flight engineer and the pilot. All members of the crew, then, interact with each other. With other types of groups the interrelationships among the members may have quite a different structure. In an office each clerk may have a completely different function, and, as far as the work itself is concerned, may deal only with the supervisor. Thus no social interchange would be necessary among the members in achieving the common goal, yet each contributes his part.

The Autonomy of a Group as a Factor Determining Its Structure. The degree to which a group is autonomous determines its structure. Some groups operate relatively independently of other groups while some groups must work in a cooperative or coordinate fashion with others. Some groups more or less determine their goals or at least the means they will utilize in achieving their goals while others have not only their goals established for them but also their method of operation. When a group is autonomous its structure can develop at the pleasure of its members. However, when it is not autonomous its structure will in part be set by some superior group or by the types of relations that exist between it and the other groups with which it comes in contact.

For example, in a factory a department that manufactures a complete item may be able to develop its own methods of production. The supervisor, together with the other members of the department, may decide that highest production is best achieved by small working teams or they may decide to specialize the activities. This, of course, presumes that higher management does not care how the work is accomplished provided production schedules are maintained and costs are minimized. On the other hand, the structure that exists among a group of workers might be imposed upon it. Management may specify in some detail the func-

tions each person is to perform and assign the duties and responsibilities. Thus a particular structure is forced upon the group.

The Homogeneity of the Members of a Group as a Factor Determining Its Structure. The characteristics of the members of the group also determine the particular structure that may characterize the interpersonal relations. If the group is quite homogeneous with respect to factors such as age, skill, attitudes, etc., there may result a group which is tightly knit. When the group is heterogeneous with respect to such factors its members may form smaller units within the larger group framework, each unit being relatively homogeneous. The members of a department may form cliques in terms of their interests, positions, or sex. The structure in such a case would be quite different from that of a group whose members have the same views, goals, and functions.

The Leadership of a Group as a Factor Determining Its Structure. The type of leadership a group has is an important factor in determining group structure. A heterogeneous group that has a leader who is both strong and understanding may through his efforts be brought together into a tight, well-functioning unit. A group of workers that has an ineffective supervisor may develop unofficial leaders in order to achieve its goals. As a consequence it may develop its own structure, or it may split into conflicting cliques.

With a supervisor who tends to be authoritarian the structure of the group may be quite rigid. Each worker would perform exactly the duties assigned to him, no more or no less. With a supervisor who is more permissive the structure of the group may be more fluid, involving an easy interchange of duties and functions among the workers.

The Size of a Group as a Factor Determining Its Structure. The size of a group is another factor that determines its structure. With a small working group of five or ten persons the structure might be quite loose, the members performing a variety of functions and changing and interchanging them readily. Communication among the members of a small group is likely to be greater, and such structure as exists is likely to develop informally rather than as a result of formal processes. On the other hand, with a large group the structure is likely to be more formal and organized. Sheer numbers preclude easy intercommunication. In order to achieve the common goal different functions have to be assigned to different persons in some prescribed fashion.

DESCRIPTION AND MEASUREMENT OF THE STRUCTURE OF WORK GROUPS

In order to understand how a work group functions as a group, it is necessary to know something about its structure. As has been pointed out, the structure of a group is given in terms of the relationships exist-

ing among its members. These relationships may be formally imposed or may arise out of informally developed interactions among the members. The relationships among individuals comprising a work group are manifold and can be said to vary in kind and in pattern. Consideration will be given here to the kinds of variables that are relevant to the measurement of group structure and to ways of indicating structure.

Kinds of Interactions among Members of a Work Group. The interactions among the members of a work group are of two kinds, interactions arising out of the work, and personal interactions. As members of an operating organization the work activities of each individual bear directly upon those of others. But in addition, as social creatures, workers develop various personal relationships as a result of being thrown together as members of a group.

Work interactions are of various sorts but two major types can be differentiated, those arising out of superior-subordinate relationships and those arising out of coordinate relationships. Since almost all business and industrial organizations are founded on the basis of vertical lines of authority, some workers always stand in superior positions to others. A foreman is the superior officer of the line worker, and the foreman in turn is the subordinate of a department manager, etc. Such formal superior-subordinate relationships are established for purposes of allocation of authority and responsibility. In addition there are certain informal superior-subordinate relationships that are established. For example, it is common to find that "old hands" are given special status by their fellows even though all are of the same rank. As a result of their age, experience, skill, or mere tenure, the advice of these "old hands" is sought by newer workers. Because of their status these "old hands" may even informally assume authority over their fellows, giving orders, making work assignments, etc. These superior-subordinate relationships constitute the problem of leadership, a topic that will be discussed later.

Coordinate relationships refer to the relationships among individuals who hold jobs of the same rank in an organization and whose work is of such a nature that the productive activities of one person influence those of others. The speed with which the fruit sorter works conditions the speed with which the packer can operate. Similarly the activities of the riveter and bucker-upper are closely integrated.

In addition to interactions that arise out of the work itself, personal relationships develop as a result of social contacts among individuals at their place of work. These personal relationships are of various kinds and are manifested in many ways. It is possible to place these personal relationships on a continuum ranging from friendly through indifferent to antagonistic. Two or more workers might find one another to their liking, and as a consequence an affinity may develop between them. These per-

sonal relations carry with them a fellow-feeling so that perhaps the individuals concerned may find pleasure in companionship off the job as well as on it. These friendly personal relationships take many forms. The workers involved may simply describe themselves as friends, they may converse with each other more frequently than with other workers, they may help each other in their work, they may have lunch and take their breaks from work together, and they may play games together.

Contrasted with these friendly personal relations are the antagonisms that may develop among certain members of a work group. These unfriendly relations sometimes come about as a result of conflicts arising out of work. In other instances they may be due to personal differences of one kind or another. In extreme form antagonisms are expressed in truly aggressive behavior. As a result of social pressure, however, they are likely to be displayed in verbal behavior as in arguments, cutting remarks, scapegoating, etc.

Personal relations need not show a high degree of consistency. Two particular persons may at one time exhibit friendly relations and at another time antagonisms. Similarly, two workers who readily help each other and who converse frequently on the job may describe themselves merely as acquaintances. Finally one worker may describe his relations with a second worker as friendly, while the second may describe his relations with the first as antagonistic. Obviously, therefore, a description of the personal relations between workers would have to be specified in time, in content, and in direction.

Quantitative Description of Interrelationships among Members of Work Groups. Recognizing the two types of interactions occurring among workers, those stemming from work relationships and those from personal relationships, it is possible to develop measures or indices of each. By noting the pattern of the interactions existing among the members of a group, the group structure can thereby be described.

It is apparent that both work and personal relationships have many aspects or manifestations. Therefore neither type can be fully described by a single kind of behavior. Descriptions of the structure of a group must necessarily be specific, referring to a particular kind of behavior related either to work or personal relationships.[8,13,14,25] More complete descriptions of group structure will involve a wide variety of different types of relationships. In some instances interpersonal relationships are established on the basis of reports by the individuals concerned, and in others on the basis of observation of actual behavior. It will be impossible here to list all of the specific kinds of behavior that have been used as indices of interpersonal relationships. A few examples will have to suffice.

The determination of superior-subordinate relationships can be taken

to illustrate the manner in which work relationships can be established among individuals. Observations can be made of the number of times each worker in a group seeks direction and counsel from every other worker concerning work problems, or each worker can be asked who he believes is the best person in his group to go to for direction and counsel about work problems. By both methods determinations could be made of the way in which workers in a group are placed in a superior position by their fellows. Superior-subordinate relations obviously could be studied by many other methods. Observations can be made with respect

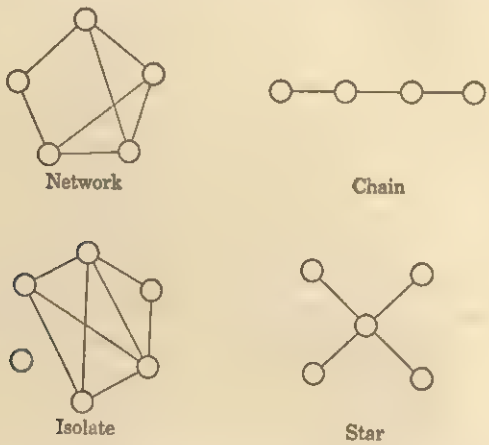


FIG. 15-2. Some typical psychometric patterns.

to which members of a group tend to give directions or orders, which members are addressed with respect as with the title *Mister*, and which members are turned to in emergencies.

Similar approaches can be utilized to ascertain the personal relationships existing among the members of a work group. Observations can be made of the number of times each worker converses with every other worker. Analysis of the nature of the remarks will indicate whether the relations are friendly or unfriendly. Questions can be asked the workers concerning which workers they like best and which they like least.

The relationships among individuals, therefore, can be stated in terms of kind, direction, and amount. The determination of these relations is termed sociometric analysis. For example, it could be said that with regard to work relationships (kind) worker A seeks help from worker B (direction) on the average of five times per day (amount), whereas worker B never seeks help from worker A. As another illustration it might be that workers C and D converse more often with each other than with any other worker, and the topics of conversation are largely personal and friendly. Such relationships are often represented graphi-

cally as in Fig. 15-2. Figures of this sort representing the interpersonal relationships are called sociometric patterns and graphically portray the structure of a group.

Some Typical Sociometric Patterns. It is obvious that the kinds of structures that groups might have are many and varied. Some typical sociometric patterns representing some focal kinds of structures are shown in Fig. 15-2. There may be a complex *network* of relations which includes all members of a group. Sometimes one or more persons are out of the group and therefore are *isolates*. *Chain* patterns are likely to occur where a hierarchy of superior-subordinate relationships has been well established. *Star* patterns are found where the activities revolve around one member who is either the formal or informal leader.

INDIVIDUAL ROLES WITHIN A GROUP

Every member of a group participates in it in different ways. The part played by some members is important in terms of achieving the group goals while that of others is quite unimportant. Some individuals are active participants while others are mere passive observers. As a member of a work group the individual is assigned certain duties or activities, he is given a certain position or status, and he assumes certain functions. These duties, positions, and functions may be termed his role in the group. Viewing an individual's place in a group from the point of view of his roles in it is helpful in understanding some of the important factors determining his behavior.[27]

Within one and the same group an individual will have a number of roles. He may see himself as fulfilling the role of the group leader, yet the role he assumes in terms of his actual behavior might be that of follower. The role played by a foreman may be that of autocrat for one worker and "father-confessor" for another. Furthermore, the roles an individual plays in a group may change from time to time. The apprentice who is upgraded to journeyman may no longer be the object of jokes. The "clown" in a working crew may become the informal leader as a result of some contribution he makes to the group during an emergency.

Kinds of Roles Assumed by Members of a Group. There clearly are many kinds of roles an individual can assume as a member of a group. Some roles are adopted for specific situations and at particular times, while others have more generality and longer duration. Benne and Sheats have suggested a three-fold classification of roles taken by members of a work group as follows: group-task roles, group-building and maintenance roles, and individual roles.[1] Such a classification is helpful in understanding the kinds and variety of roles that are adopted.

Group-task roles are related to the achievement of the purposes of the group. They are concerned with the facilitation and coordination of the group's effort, and with the definition and solution of the problems facing the group in connection with its purposes. Many different functions are subsumed under this category. Following are some examples. The "energizer" is one who attempts to prod the group to greater activity. The "orienter" tries to define the position of the group with respect to achievement of its goals, summarizing progress that has been made, pointing out deviations from goals, etc. The "opinion seeker" is one who is concerned with a clarification of values with respect to the group goals.

Group-building and maintenance roles are concerned with the functioning of the group as a group. They are oriented toward maintaining or altering the group way of working so as to regulate and perpetuate the group. Examples of roles of this kind are the "encourager" who praises the contributions of others, and the "compromiser" who seeks to obtain concessions from disputing members of the group so that its existence will not be endangered.

Individual roles refer to those that are directed toward satisfaction of the individual's needs. Their object is to facilitate the achievement of some individual goal that has no particular relevance either for the group task or for the functioning of the group as a group. They are manifest in group behavior merely because the group is a means for facilitating or impeding the achievement of personal goals by the individual. Such roles are seen in the "playboy," the "aggressor," and the "help-seeker."

Roles as Stereotypes. Frequently the roles assumed by an individual or informally assigned to him by other members of his group are stereotypes. A stereotype is a belief that is simplified in content and is likely not to correspond with the facts of the matter.[17] Stereotypes are beliefs widely held in a particular culture group. Thus among a group of workers the "big boss" may be thought of as a person who is cold and distant. He may actually be warm and friendly. Similarly all apprentices as a class may be regarded by older workers as flighty and undependable, whereas they actually may be serious students of their trade.

Factors Determining an Individual's Roles in a Group. A worker's roles in a group are determined by three sets of factors. First of all, it is obvious that functions are formally assigned to the worker by the group of which he is a member. Secondly, a role may be assumed by an individual as a consequence of his unique personality. Finally, and less frequently recognized, is that fact that roles are thrust upon a worker by the group or groups of which he is a member. Depending upon the nature of the group and the characteristics of the individual, one or two of these factors may be more important than the others. There will seldom be a

situation wherein the roles of a worker are not determined to some degree by all three factors.

The roles formally assigned a worker by the group of which he is a member may result from the position he holds and from recognition by the group of his particular traits and abilities. By virtue of the position he occupies, a worker is expected to perform certain functions. An individual assigned to the job of inspector is expected to examine manufactured articles for deficiencies. The formal role given him is that of judge of the work of others.

In carrying out a formally assigned role, even in a well-established organization in which job functions are specifically defined, there is still some latitude for individual differences. A particular job might be performed equally well by two individuals who operate in different ways. Some organizations, therefore, may seek to take advantage of the personal attributes of the individual and adjust the functional characteristics of the job accordingly, that is, change the formal role associated with it to take advantage of the particular strengths of the individual who is assigned to it. Thus while the function of department heads may be stated generally as supervisory, for one individual management may formally specify the role as "expediter," and for another individual as "organizer," thereby seeking to utilize more fully the best talents of each.

Recognizing that the specifications for jobs are seldom so rigidly laid down but what there is some scope for individual differences, it is not surprising to find that the roles a worker takes in his organization are determined to some degree by his unique and individual personality. Two workers on the same job who possess different patterns of traits and abilities may assume very dissimilar roles in their organization. One foreman with a highly developed feeling for authority and responsibility may adopt the role of the "distant but fair boss," while another who is a sociable person may seek to establish himself as the "buddy of his men." An individual projects into his job aspects of his personality which are reflected in terms of the role he assumes relative to the other members of the organization.

Finally, roles may be imposed informally upon the worker by other members of the group. Sometimes the roles result from the reactions of his fellows to him as an individual and sometimes from the particular position in the organization to which he is formally assigned. An older worker in a group may have the role of "fatherly adviser" thrust upon him because of the wisdom expected as a result of his age. A worker better educated than the others in a group may be considered a "dude." In certain situations particular roles appear to be associated with particular jobs. Roles are imposed upon persons in certain positions by virtue of the positions themselves rather than as a result of the characteristics of the

individuals who happen to be occupying them. Thus inspectors sometimes are considered spies of management because they are inspectors, and supply clerks are regarded as being fussy old maids because they are supply clerks.

Maintenance of Roles. If an individual's environment did not possess a considerable degree of constancy, the world for him would be confusing and unpredictable, and his adjustment would be exceedingly difficult. The individual, therefore, not only imparts to his physical environment but also to his social environment a considerable degree of constancy.[17] In the physical world, for example, a familiar object a long distance away is not seen as being smaller than when it is close, but rather it is seen as being approximately the same size under all circumstances.[4] When a horse is seen at a distance it is not perceived as a pony or colt. Similarly, a worker who considers his foreman to be a kindly and fair person will tend to maintain this belief regardless of the foreman's actions.

In the context of roles it could be said that the role one individual attributes to another influences his perception of the behavior of the other, bringing that behavior into conformity with the role. The worker who casts his foreman in the role of a kindly and fair person is unlikely to accept as true any actions of his foreman that do not conform with that role. The worker's preconceptions about his foreman influence the ways in which he interprets his foreman's actions. Hence actions on the part of the foreman that others would deem unjust would be rejected by the worker as not having occurred, or would be explained away by him in his endeavor to maintain constancy in this aspect of his social environment.

There are, of course, circumstances wherein an individual will modify his conceptions of things or people. One such circumstance is when the concept is not firmly established.[17] For example, a worker may not have had time to develop very clear ideas about his foreman and may tentatively be regarding him as being kindly and fair. One action by the foreman that the worker could consider unjust may then change his perception of the foreman, with the result that he would now cast him in the role of a dishonest and cruel person.

Facts that are inconsistent with a role that is firmly established, as in the instance of stereotypes, will in one way or another tend to be rejected by the individual. In this way he maintains constancy in his social environment. Ways in which stereotypes are maintained can be illustrated from the results of a study by Haire and Grunes.[10] These investigators presented college students with a list of "facts" about a workingman, and asked them to describe the sort of person they thought the facts depicted. The "facts" were such items as the following: he works in

a factory, reads newspapers, goes to the movies, and cracks jokes. The resulting descriptions fitted the pattern of a "typical American Joe," a healthy, happy, uncomplicated, mildly sociable person who, while not too bright, tries to keep abreast of current trends and is interested in simple pleasures and undistinguished activities. To another group of students given the same instructions, the list of "facts" was increased to include the item "he is intelligent." In seeking to maintain their stereotype of a workingman these students either denied that the man was intelligent, distorted the term intelligence to give it a different meaning, or denied the man was a workingman. Thus in one way or another the incongruous fact was explained away and the stereotype was maintained.

This tendency to maintain the characteristics of roles assigned to others has important results in the industrial situation. If the workers in an organization view the management as being hostile, friendly or even innocuous actions on the part of management may be interpreted as being counter to their welfare. Similarly, if management views workers as being clock-punchers who are interested only in their wages and not in their jobs, then when workers do more than is expected of them such activities may be viewed with suspicion.

Roles as Prescribed Behavior and as Perceptions. Roles can be thought of as expectancies, that is, anticipations that an individual will behave in certain ways and not in others. As Newcomb has pointed out, a role has certain prescribed and certain proscribed aspects.[24] The position that any individual occupies in an industrial organization or in an informal group carries with it a commitment to behave in a particular fashion. Such behavior is conceived as being a necessary part of the individual's role. On the other hand, there are certain things that an individual in a particular position should not do, behavior that is proscribed or forbidden. Finally, there are certain things that the individual may or may not do, behavior that is permitted but not demanded.

In a particular organization, for example, the role of the foreman may be that of "encourager." He would then be expected to take actions that stimulate his subordinates toward greater achievement, and to avoid negative criticism of their work. Supportive actions are prescribed for this role and condemnation is proscribed. He may or may not eat lunch with them, he may or may not encourage them to participate in planning, activities, etc. Such actions have no direct bearing upon the role of "encourager" and hence they are permitted but are just not pertinent in delimiting it.

The necessary behavior for a particular role may be prescribed by the individual filling it, by the members of the organization, or by the informal group to which he belongs. From various formal and informal sources an individual develops certain notions concerning the role that

is expected of him. He then develops beliefs concerning the ways in which he is expected to behave. Such roles may be termed self-prescribed roles. Thus the dispatcher in a taxicab company may view his position as being that of director and coordinator of the activities of the driver, and statuswise as being higher than the job of driver.

The organization or informal group may prescribe the role to be associated with a particular position, expecting certain functions, status, and activities of the individual filling it. The role of an individual filling a certain position may be prescribed by his group. When a group of work-

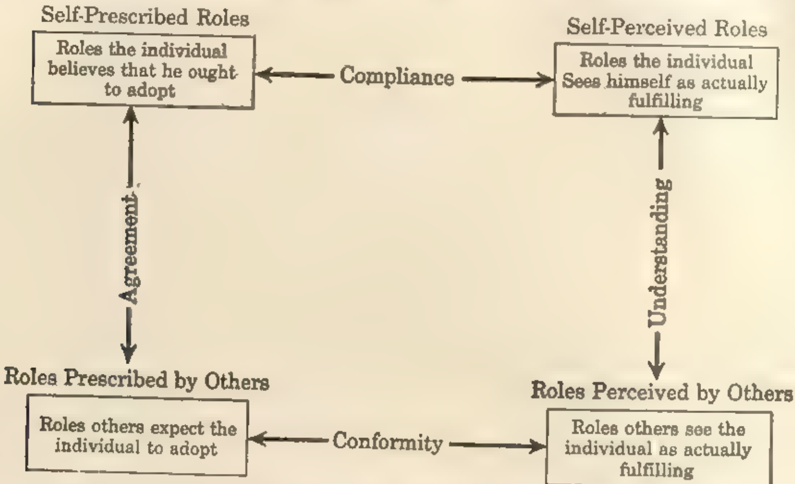


Fig. 15-3. The relationships between prescribed and perceived roles.

ers informally assigns the role of "clown" to one of their fellows they expect him to jest and to make fun of any situation that occurs because for them that is the way a "clown" behaves. Similarly, if workers believe that management is unfair, they may interpret any action taken by their foreman as unjust no matter how considerate he may try to be.

There are self-prescribed roles—roles the individual believes that he ought to adopt, and roles prescribed by others—roles others believe that the individual should adopt. However, the individual may not adopt the role he thinks he should, and others may not believe he is adopting the role they think he should. That is, the self-prescribed role may not correspond with the role the individual seeks to assume and perceives himself as fulfilling. Similarly, the role others prescribe for the individual may not correspond with the role that they perceive him to be actually fulfilling. This four-fold categorization of roles is shown graphically in Fig. 15-3.

Lack of correspondence between prescribed roles, between perceived roles, and between prescribed and perceived roles produces situations

conducive to friction among individuals and among groups. When people behave in ways that are neither desired nor expected by themselves or by others, interpersonal difficulties are likely to develop. Therefore examination of the degree of correspondence between roles will be fruitful in terms of understanding some of the determinants of intragroup and intergroup conflicts.

Correspondence between Self-prescribed Roles and Roles Prescribed by Others. The correspondence between self-prescribed roles and roles prescribed by others is a matter of degree of agreement on the character and scope of the individual's position in the organization or informal group. In terms of formal roles the concern is with the job specifications. In the formal views of an organization an individual holding a given job is obligated to perform certain prescribed functions, to have a particular status, and to engage in specified activities. However, for one reason or another the worker holding the job may not have the same views. In such a case his self-prescribed role and the role prescribed by others for him will not be congruent.

Disagreement may arise as a result of poor definition of the job or as a result of poor communication. In many instances the duties and responsibilities allocated to a job are not clearly described. A salesman who is told that it is his responsibility to follow up on orders may not check whether the merchandise he sells is actually shipped because he does not interpret "to follow up" as having this meaning. For him to follow up means to keep in touch with customers. Even though the job specifications are clear the worker might not have this information communicated to him in a manner that is understandable to him. This situation is likely to occur when the job specifications are not formally recorded but are simply known to the superior. By word of mouth the superior must convey to his subordinate the duties and responsibilities connected with his job. When communication of the information is unclear or incomplete, then the role that the worker himself prescribes for the job will differ from the role other members of the organization prescribe for it.

Disagreement with respect to prescribed roles may also result from differences in objectives, or differences in interpretation of objectives. The salesmen in an organization may see their role as consultants to the clients and their function as building up substantial and continuing accounts. Management, on the other hand, may consider the job of salesman as one of discovering and developing new accounts, with the maintenance of the accounts a function of service personnel.

A role informally prescribed by a group for its members may be at variance with that which the individual prescribes for himself. In order to distinguish themselves from those in lower jobs, the senior clerks in

an office may traditionally "go out for lunch" rather than carry it to work in a paper bag. A new senior clerk brought into the organization from another establishment might have worked in a situation in which carrying one's lunch was the tradition of all salaried employees. The role that he prescribes for himself would therefore permit of lunch carrying, a behavior proscribed by the new group in which he finds himself.

Correspondence between Self-perceived Roles and Roles Perceived by Others. The functions, status, and activities an individual perceives himself fulfilling may not be those that others perceive him to be doing. A foreman who thinks of himself as being fair and impartial in dealing with his subordinates may be regarded by them as being strict and biased.

The correspondence between self-perceived roles and those roles as perceived by others involves the problem of mutual understanding.[18] The problem here is the degree to which others understand, comprehend, and have insight into the role that a particular individual believes himself to be fulfilling. The foreman who wishes to be fair and impartial takes actions that he believes will implement his objectives. His subordinates, however, may misinterpret these actions and fail to understand his objectives.

Lack of correspondence between self-perceived roles and those roles as others perceive them does not result from a disagreement concerning the facts of a matter but from a disagreement with respect to the interpretation of a set of facts. In the case of the foreman just mentioned his subordinates may complain that he is cold and aloof. The very behavior of the foreman that leads his subordinates to this conclusion may be deliberate on his part but for the different objective of effecting impartiality. Both parties would agree on what specific behavior was manifest, but the subordinates would consider that it results from one motivation while in fact it results from a very different motivation on the part of the foreman. The subordinates therefore fail to understand the true motivation underlying the foreman's behavior, and he in turn fails to understand how his behavior can be misinterpreted by his subordinates.

Correspondence between Self-prescribed Roles and Self-perceived Roles. When an individual prescribes a role for himself he is, in effect, formulating for himself a concept concerning his obligations. In view of all of the information he possesses about his job and his place in the working group, he arrives at a set of ideas relative to the duties and functions he should perform and the position he should assume in the group. Until this conception is challenged by others, this role that the individual constructs is, as he sees it, his appropriate role, the role that he believes the group expects of him.

For one reason or another the individual may decide to reject this self-prescribed role either in whole or in part. He is unwilling to conform

to the role he believes is expected of him. As a consequence he behaves in a different way, assuming a role which he perceives is at variance with the prescribed one. The self-prescribed role will then not correspond with the self-perceived role—the role he sees himself as actually fulfilling. It is apparent that the degree of correspondence between self-prescribed and self-perceived roles is a matter of compliance. It concerns the individual's willingness to accept a set of obligations as he believes them to exist. He will not yield to the demands that he believes are being made of him.

From the individual's point of view, when the role he sees himself as actually fulfilling differs from the role he prescribes for himself, he is violating rules or customs. This would in fact be true if his self-prescribed role agreed closely with the group-prescribed role. However, if his self-perceived role agreed with the group-prescribed role he actually would be conforming to rules and customs, and any conflict that he feels as a result of not fulfilling what he believes his obligations to be would be false.

Suppose that in the formal organizational structure of a factory the safety engineer is one of several specialists who are directly responsible to the plant manager and are placed in staff positions rather than in executive or supervisory positions. As a result of this situation the safety engineer may consider that his activities should largely be advisory, involving the study of safety problems in the factory and making suggestions for action to the plant manager. In constructing a picture of what he thinks his job out to be, he would prescribe for it only advisory functions directed toward his immediate superior. He would not consider it within the scope of his activities to give direct orders to the heads of departments concerning safety measures to be introduced and maintained. He would proscribe such activities for himself, believing that they are not appropriate functions of staff personnel but only of those who are in charge of some operation in the factory. However, being genuinely interested in increasing safety in the factory, the engineer may take it upon himself to issue instructions directly to operating personnel concerning safety procedures to be introduced and maintained. These activities are part of the role that he perceives himself as actually taking and clearly conflict with the role that he prescribes for himself.

Correspondence between Roles Prescribed and Roles Perceived by Others. The role that others prescribe for a given individual may be thought of as a standard they set for evaluating the behavior of a person in his position and with his characteristics. The president of a bank may expect the vice presidents to be conservative in action and opinion. A vice president who wears sport clothes at work would be considered to be conducting himself in an inappropriate manner. Old hands expect a

new worker to be self-effacing, and one who is not is considered brash. Therefore, from the point of view of others, when the role they see the individual as actually fulfilling does not correspond with that which they expect of him, the individual is judged as failing to conform with rules or customs. Correspondence between roles prescribed by others and the roles as perceived by them has to do with the problem of conformity.

The correspondence between roles prescribed and perceived by others can be considered in relation to a formal organization. In this framework the correspondence becomes the basis for superiors' evaluation of the job performance of their subordinates. The prescribed roles are the job and worker specifications and other standards of job performance. The perceived roles are the job performances of subordinates as their superiors consider them actually to be. The greater the similarity between the two roles the higher will be the superiors' appraisals of their subordinates.

For example, a department head in a factory may consider that the functions of his foremen are to transmit work orders to the men, to plan and schedule the flow of work, to keep records on attendance and output, and to handle disciplinary problems. A foreman whom the department head believes does these things would be considered by him as a good foreman. But one who comes to the department head for advice on disciplinary problems or who introduces new methods of work would not be considered as wholly satisfactory.

As pointed out earlier, when the members of a group agree closely on the characteristics of the role they expect of an individual of a given type, then the role is a stereotype. A role prescription of this sort when made by a group becomes a social norm for it. An individual who in the eyes of his group fails to adopt the role it prescribes for him is violating its customs. The failure of an individual to conform may result in his ostracism or punishment.

Thus, piece-rate workers may informally agree among themselves with respect to the amount of production they will individually turn out in a day.[8] The role of a "good guy" proscribes other than this amount. Those who turn out more are "rate-busters," and those who turn out less are "bums." Supervisors may actually support the "good guy" role in workers, condoning restriction of output in order to show uniform production in their departments. A worker who fails to conform with this role may be threatened or ridiculed by his fellows and disciplined by his supervisor.

LEADERSHIP

An organization has been defined as a group in which the members are differentiated in terms of their functions as these relate to the achievement of the common goal. One type of differentiation of function is hori-

zontal, whereby different individuals perform different functions but on the same level; that is, the functions are coordinate in nature. Thus in a fruit-processing plant some workers sort the fruit while others pack it. Another differentiation of functions is vertical, whereby some individuals guide or direct the activities of others; that is, the functions involve a superior-subordinate relationship. A foreman in the fruit-processing plant gives orders to sorters and packers alike, regulating their activities. It is from this latter type of differentiation that the problem of leadership emerges.

Definition of the Term Leadership. The concept of leadership is by no means a clear one. Certainly one of the major sources of confusion is in the difficulty of defining the term. Examination of definitions that have been evolved show that the term is used with a wide variety of meanings.[23, 29] Some of the definitions are connected with the position a person holds in an organization, others with his personal qualities, and still others with his functions.

Persons in certain higher positions in an organization are sometimes considered to display leadership by virtue of the very nature of the position they hold. Frequently a supervisor is considered to be a leader merely because he has the job title of supervisor. It is necessary, however, to distinguish between real leaders and men in high-status positions.[9] Persons placed in higher positions in a formal organization may show none of the behavior or characteristics associated with leadership, hence position alone cannot be considered the criterion of leadership.

Individuals who possess qualities that are superior to those of their fellows may be considered to be leaders simply because they possess these qualities. The qualities may be certain knowledges and skills required for the group's activities. Because of his experience, an "old hand" among a group of workers may be considered a leader. He more than others is able to deal with difficult and unusual problems, he has greater skill in adjusting and operating machines, and he has a more thorough understanding of the factors involved in any situation that arises.

The particular pattern of personality characteristics that an individual possesses may result in others considering him to be a leader. Because of his calmness during action, a fireman may be an exemplar for his group and therefore be thought of as evidencing leadership. However, it may be argued that persons of superior qualities are not necessarily leaders since there may be no deliberate attempt on their part to influence the behavior of others. Furthermore, even if there is such an intent the behavior displayed may not be termed leadership. For example, the "old hand" who shows another worker how to do something is merely aiding or instructing him rather than leading him.

It is sometimes said that leadership is the exercise of certain functions

and the definition of leadership becomes one of defining these functions. Some writers insist that leadership is manifest only when the leader motivates workers by some kind of indirect process, and others hold that leadership involves direct control over the activities of the members of an organization. Points of views such as these consider that leadership exists only when the stated functions are exercised. No other functions or activities are considered as leadership. The difficulty with definitions of leadership based upon functions is that the definition of the term is confused with a particular concept or theory of effective leadership. For instance, one who holds that democratic leadership is the only effective type of leadership then defines that type of activity as leadership and considers the direct giving of orders not as leadership but rather as some other kind of process. Therefore a definition of leadership is required that does not confuse semantics with values.

It becomes apparent that leadership cannot be defined just in terms of position in an organization, personal qualities, or functions exercised. An individual may be a leader whether or not he holds a certain position in an organization, possesses certain personal qualities, or exercises certain functions. To be sure he may be a leader because of his position, personal qualities, or functions, but they in and of themselves do not define leadership. Leadership becomes recognizable only in terms of what people do. For purposes of study, leadership should be defined in terms of action or behavior.[28]

One of the most useful definitions of leadership is that developed by Shartle, Stogdill, and a group working at Ohio State University.[31, 32] These investigators define leadership as the process by means of which the activities of the members of an organization are influenced with respect to goal setting and goal achievement. The social conditions that permit the existence and manifestation of leadership are a group, a common task, and differentiation of responsibility. Since, as pointed out earlier, these are the conditions under which a group becomes an organization, it is apparent that leadership emerges as one of the characteristics of an organization.

Defining leadership as the process of influencing the goal setting and achievement of an organization is helpful since such a definition does not impose any specific concept or theory of effective leadership. Furthermore it does not restrict leadership to individuals with particular personal qualities or in particular positions. The value of this definition is its flexibility in covering a wide variety of behaviors while at the same time restricting leadership to its proper setting—the organization.

By defining leadership as the degree to which an individual influences others in his organization toward the common goal, the focusing of attention solely on persons in supervisory, administrative, and executive posi-

tions is avoided. This is important since sometimes the true leaders in a group are not formally designated as such, but rather informally assume the leader's role or are unofficially assigned it. Leadership is manifest in different ways and to different degrees by different members within a single organization. Sometimes the leader's influence is direct and its effects easily observable. In other instances it is indirect and subtle. But it is not confined solely to those in so-called leadership positions.

Identification of Leaders and the Measurement of Leadership. From the foregoing definition of leadership it follows that a leader is anyone who influences the members of an organization with respect to goal setting and goal achievement. It is clear that those persons who hold supervisory, administrative, or executive positions in an organization cannot be considered leaders merely because they hold such positions. While the probability is that most of them will operate as leaders there is no certainty that they will. Furthermore, others who do not hold such formally designated positions may in fact function as leaders. Consequently methods other than the designation of formal positions in an organization must be used to identify leaders.

In any given organization it would not be unexpected to find that every member influences every other member with respect to goal setting and goal achievement. In such a situation each member would be manifesting leadership and hence all would be leaders. There would be differences, of course, in terms of the amount of influence, with some persons exerting more than others. Leadership therefore must be considered as varying in degree or amount among the various members of an organization. Any dichotomy of persons into leaders and nonleaders or followers is completely arbitrary. It might be argued that leaders are those who exert the greatest influence. This notion, however, presupposes that degree of influence can be measured and that a point on the scale is readily established below which degree of influence is quite unimportant. In any event it is clear that the problem is not one of discovering which persons in a group are the leaders but rather one of measuring the degree of leadership manifest by every member of the organization.

Most of the techniques developed for identifying leaders or measuring leadership approach the problem indirectly. Instead of measuring the amount of influence an individual exerts upon an organization, attention is given to other aspects that, in turn, are likely to be related to his degree of influence. A wide variety of specific techniques have been developed.[20,23] Among the most widely used are rating methods whereby an individual's ability as a leader is judged by his superiors, his peers ("buddy" ratings), or his subordinates. One commonly used rating method is the so-called nominating technique.[6,26,34] In this procedure persons acquainted with other members of an organization name the one, two, or

three individuals in it they consider the best leaders, and perhaps the one, two, or three they consider the poorest. It is presumed that the higher an individual is rated, or the larger the number of positive nominations he receives, the more likely he is to have influenced group behavior. While this procedure has many practical advantages its validity has yet to be adequately examined.

Several techniques for measuring leadership involve determinations of the amount of time the members of an organization spend in "leader" activities such as supervising, planning, evaluating, etc.[28] This type of approach makes considerable sense since the activities studied are those ordinarily expected of supervisors, administrators, and executives. However, while this approach might be useful in studying those in formally designated leadership positions, these particular activities may not apply equally well to informally designated leaders. Furthermore, doubt may be cast on the reliability with which time estimations are made and whether the mere performance of the activities themselves actually has any effects upon the members of a group in respect to goal setting and goal achievement.

The Characteristics of Leadership Ability. Many discussions of leadership imply that the capacity to influence others is a single unitary trait possessed by different people in different degrees. The notion is that those who possess a great deal of the trait are necessarily successful leaders while those who possess very little of it are poor leaders. Reviews of the findings of investigations of the personal characteristics of leaders throw considerable doubt on the notion that there is a consistent general trait of leadership.[15,30]

If leadership were a general capacity, then it would be expected that every individual who is a good leader in one situation would be an equally good leader in all other situations, and that every individual who is a poor leader in one situation would be an equally poor leader in all other situations. In other words, if the same group of individuals were placed in a number of different situations and their effectiveness as leaders was measured in each, the coefficients of correlation among the various measures of leadership effectiveness would be quite high. Pertinent investigations of this problem indicate that such correlations are in fact quite low.[5,7] It has been estimated that under the best of circumstances and utilizing highly reliable measures of leadership ability, the correlation between performance as a leader in different situations would be no greater than .35.[20]

Because the correlations are so low it is apparent that if such a general trait does exist it is not of great importance. While those persons who are good leaders in one situation will tend to be good leaders in another and those who are poor leaders in one situation will tend to be poor in

another, there nevertheless will be a large proportion of reversals. Thus it is likely, but far from certain, that an individual who has been a good supervisor of office workers will also be a good supervisor of sales personnel.

If in some particular situation the proficiency of persons as leaders is measured in several different ways, as with different types of ratings of leadership, the coefficients of correlation among such measures will be found to be reasonably high.[7, 20] Within any given situation an individual's proficiency as a leader as indicated in various ways shows a substantial degree of consistency. These coefficients of correlation between different measures of leadership ability in one situation are considerably higher than the coefficients of correlation between measures of leadership ability obtained in different situations. The consistency of leadership behavior within any particular situation, then, is greater than the consistency between different situations.

It is apparent that proficiency as a leader is to a considerable extent situationally determined.[11] In one type of social situation a particular individual might perform well as a leader while in another he might perform poorly. For a second individual the reverse might be true. Consider the case of the individual who has little tolerance for ambiguity. Such a person needs to operate in a framework of clearly defined functions, rules, and policies. If the organization of which he is a member is highly structured his performance as a leader is likely to be excellent. However, if the organizational structure is pliant and changing he may be unable to operate at all as a leader. On the other hand, a person who has a great need for individual action would find the highly structured organization too constricting and would operate much better as a leader in one that is more flexible.

Therefore it can be concluded that there is no single ability of leadership but rather a number of different abilities that are loosely grouped together and called leadership ability. If there is a general trait of leadership that plays a part in all situations it is relatively unimportant in determining an individual's success as a leader. To a considerable extent the manifestation of leadership is determined by the social situation. Under one set of circumstances an individual will be a good leader and under others he will be a poor one.

The Functions of the Leader. Since leadership is to such a large extent situationally determined, there is certainly no single set of functions that are performed by all leaders. In one situation the leader will perform one set of functions and in another situation his functions may be quite different. The functions a leader performs vary with a number of different factors. The relative importance of different leadership functions depends in part upon the type of organization in which the leader operates, the

particular social situation existing within the organization, the nature and size of the group membership, the personal characteristics of the leader, and whether he is a formally designated leader or only informally assumes the leader's role.

Krech and Crutchfield have attempted to describe the various functions performed by leaders.[17] All leaders will of course not perform all of these functions, or at least not to the same degree. Hence the list can be thought of as portraying the kinds of functio characteristic of leaders in general rather than the functions of any given leader. The 14 functions listed by Krech and Crutchfield fall into three categories: functions bearing directly upon the setting and achieving of organizational goals, functions bearing upon the operation of the organization, and functions relating to the leader as a figure for the group.

Functions of the Leader in the Setting and Achieving of Organizational Goals. In connection with his activities bearing directly upon the setting and achieving of organizational goals, perhaps the most obvious function of the leader is that of executive. In his role as executive he is responsible for seeing that the appropriate activities of the organization are carried out. A foreman assigns tasks to workers and sees that these tasks are properly executed.

Another function performed by the leader is policy making. He may either establish organizational goals and objectives himself, or he may participate with his superiors or subordinates in establishing them. Thus the president of a company, perhaps in collaboration with the board of directors or with his staff, determines the nature of the commodities or services with which the company will be concerned.

A final function, planning, is intermediate between the determination of policies and their execution. In this connection the leader makes decisions concerning the ways and means the organizational goals can be achieved. A foreman not only assigns tasks to his subordinates but he may also plan work schedules and devise operational procedures.

Functions of the Leader in the Operations of the Organization. The second group of functions performed by the leader is connected with the operations of his organization. Six functions can be differentiated in this category. First, the leader is an expert in the principal activities of the organization. Frequently supervisors are promoted from line positions primarily because of the proficiency they displayed on the lower job. The technical information and skills supervisors possess are useful in aiding and instructing their subordinates in effective work procedures.

A second function of the leader in this general category is as external group representative. There are many reasons why it may be necessary or desirable to have a single representative of an organization deal with outside individuals or groups. The organization may be too large to

function as an entity in this activity because taking a large number of workers from their duties may interrupt work schedules and thus be uneconomical, agreements may be more easily achieved by individual than group action, etc.

A third function of the leader is as a surrogate for individual responsibility. The leader relieves other members of the group of certain responsibilities, and they in turn place their trust in his decisions. In an informal group of workers, one individual may be given the responsibility of conveying complaints to the supervisor. The others trust him to present their points of view and accept any approach to the supervisor he may wish to make.

A fourth function of the leader is as controller of internal relationships within the organization. Thus it is the task of the manager of a plant to see that the various departments in it coordinate their activities.

A fifth function of the leader is as administrator of rewards and punishments. As leaders, supervisors encourage, upgrade, and promote deserving workers, and reprove, transfer, and fire poor workers. Even leaders in informal groups concern themselves with discipline. The leader of such a group may call upon its membership to oust an errant member.

Finally the leader acts as an arbitrator and mediator, seeking to maintain harmony among the members of the organization. The president of an organization tries to keep the peace among competitive and ambitious vice presidents. The leader of an informal group tries to prevent it from breaking up into opposing factions.

The Leader as a Figure for the Group. The third group of functions of the leader, those connected with his position as a figure for the group, are less well recognized than the other two groups. The leader's activities and influences in this connection are likely to be so subtle as to be overlooked. Five functions can be differentiated in this category. In many organizations the leader serves as an exemplar, a model for others to simulate. The military leader who precedes his troops into battle, the office manager who invariably is 15 min. early at his desk in the morning, the foreman of a crew of power-linemen who always checks his safety equipment before climbing a pole, are all seeking to influence the members of their groups by being good examples.

A second function in this group involves the leader as the symbol of the group. In this role he provides a kind of continuity and stability for the group, standing for it despite changes in circumstances and membership. Thus the chairman of the board of a company who may in fact exercise little or no control over it may in the minds of the stockholders, the workers, and the public stand as a visible sign of the purposes, the spirit, and the very properties of the company.

Thirdly, the leader functions as an ideologist. By presentation of his ideas concerning the group he is a source of beliefs about it. The official ideology as given by the leader may or may not reflect the ideas of the group. For example, the designated spokesman for an informal group of workers may convey to the supervisor altogether incorrect notions with respect to the complaints and dissatisfactions of the workers.

Fourthly, the leader may function as a father figure, fulfilling an emotional role for the members of the group. By identifying themselves with their leader, the members of a group draw strength and feelings of security. Statements from workers such as "The old man will always be behind us" typify this situation.

Finally, the leader functions as a scapegoat. He provides a ready target for the aggressions of the members of the group. Failure can be projected upon him. A department head in a plant may attribute the inadequate production in his department to insufficient cooperation on the part of the plant manager.

Types of Leadership. The varieties of leadership are so extensive that attempts to classify its different manifestations into types is extremely difficult. There is some agreement in distinguishing two types, autocratic leadership and democratic leadership. However, as Krech and Crutchfield have pointed out, the use of such terms to describe leadership situations is largely one of convenience.[17] The danger lies in the possibility that more may be ascribed to a particular type of leadership than is warranted because of the inferential meanings drawn from the particular term attached to it. Furthermore, it is doubtful if in actual practice either of these two types of leadership actually exist in pure form. They are to be thought of more as reference points or extremes rather than as existing kinds of leadership.

With authoritarian leadership the leader alone determines policy and makes plans. Only the leader dictates the activities of the members, gives orders, and passes information along to the other members of the organization. As a consequence the leader sets the pattern of the interrelationships among the members. In a case of authoritarian leadership the sociometric pattern is a star, with relationships existing only between the leader and the members but not among the members. The leader is the key person in authoritarian leadership. The whole operation of the organization depends upon him. In his absence it may function inadequately or not at all.

When democratic leadership occurs the entire group is involved in and accepts responsibility for goal-setting and achievement. Therefore, in democratic leadership the leader's activities are quite different from those in authoritarian leadership. Through group discussion the democratic leader attempts to get every member to participate to the maxi-

mum in the group's activities and to spread responsibility widely among members. Part of the leader's task is to encourage and reinforce constructive interrelationships among members and to reduce intragroup conflict and tensions. The sociometric pattern for democratic leadership is a network which involves a tight pattern of complete interrelationships among all members. While the leader is quite an important figure in a democratic situation, he is not the key figure that he is in an authoritarian situation. He serves more as a coordinator or agent for the group. Hence the group is not dependent upon him as an individual and can function effectively in his absence.

The Relative Effectiveness of Authoritarian and Democratic Leadership. Numerous studies have been conducted comparing the relative effectiveness of authoritarian and democratic leadership.[16,19,22,33] These studies have considered a wide variety of individuals and groups, and have evaluated the effectiveness of leadership in many different ways including group productivity, morale, and satisfaction of the individual members. The findings are quite consistent in pointing up the general superiority of democratic over authoritarian leadership.

As a consequence of these studies the definite conclusion is often reached that for effective operation an organization must have democratic leadership. While there may be no reason to doubt the facts found in comparative studies of authoritarian and democratic leadership, there are, as Krech and Crutchfield have indicated, certain limitations in the generalizations that can be drawn from them.[17] First of all it should be noted that these studies are ordinarily conducted in a democratic culture. It is entirely possible that in an authoritarian culture, or in an organization that has an authoritarian tradition, democratic leadership may be less effective than authoritarian leadership. Secondly, some group members do function quite effectively under authoritarian leadership and some function quite ineffectively under democratic leadership. Persons who are submissive or whose social background is authoritarian may be quite satisfied with an authoritarian situation. Finally, authoritarian leadership is supportive since it relieves the individual of responsibility and decision making. Authoritarian leadership, therefore, would be expected to be more effective than democratic leadership when the members of a group are emotionally insecure or find themselves in an ambiguous or critical social situation.

While the foregoing limitations on the interpretation of the findings of leadership studies should be borne in mind, the importance of the general trend of the findings indicating the superiority of democratic leadership should not be minimized. It is apparent that under most circumstances in this country and with most individuals and groups, democratic leadership is more effective than authoritarian leadership. The

tradition in this country is democratic. Individuality rather than submissiveness is stressed as a desirable trait. Efforts are made to minimize situations leading to feelings of insecurity, thus reducing or eliminating the need for the strong supportive leadership of the authoritarian kind.

The Conditions of Effective Leadership. To say that effective leadership is democratic in nature is too general a description to be of practical use. What is needed is a more detailed specification of the conditions under which leadership is effective. Many studies have been conducted with the purpose of specifying such conditions and a wealth of information is therefore available. Different investigators have employed so many different terms in describing their results, however, that interpretations are very difficult to make.

An analysis of the conditions of effective leadership presented by McGregor provides one of the best integrations of the experimental findings.[21] Since McGregor views leadership in terms of the relationships between superiors and subordinates, his analysis is most pertinent to those situations where leadership resides in those in formally established leader positions such as executives, supervisors, and foremen. However, the conclusions apply almost equally well in cases of informal leadership, and especially when the informal group is highly structured.

According to McGregor, the outstanding characteristic of the relationship between superiors and subordinates is the dependence of the latter on the former. The subordinate is dependent upon his superior for the satisfaction of many of his needs. The subordinate depends upon his superior for continuity of employment, promotion, increased pay, and a variety of social satisfactions. Even when the subordinate is afforded formal protection from arbitrary actions by his superior by formal contract or by some informal circumstance, the dependent relationship still exists. A foreman can "throw the book" at a worker who is protected by a union. He can in many subtle though legal ways make the subordinate's life quite unpleasant. The president of a company can summarily dismiss a vice president on the basis of a flimsy rationalization.

The dependent nature of the relationship between superiors and subordinates is the same regardless of the level of the individuals in the organization. The line worker is dependent upon his foreman, the foreman upon his department supervisor, the department supervisor upon his division superintendent, the division superintendent upon the plant manager, the plant manager upon the vice president for operations, the vice president upon the president of the company, the president upon the board of directors, and the board of directors upon the stockholders. The dependent relationship between any given subordinate and his superior, therefore, is merely one link in a chain of dependent rela-

tionships. Except at the extremes, each individual has two roles, one as subordinate and one as superior.

As a consequence of the dependence of the subordinate upon his superior, the subordinate has a need for security and a need for independence. First of all the subordinate seeks to secure his position in the work situation. He struggles to protect himself from any threats, real or imagined. Guarantees by management or by union contract provide some security but since they are concerned largely with such matters as wages, retirement plans, and health and accident insurance, they do not get at the heart of the matter—the personal dependence of the subordinate upon his superior.

When the subordinate has achieved at least a reasonable degree of security then his need for independence becomes manifest. He seeks to overcome his dependence upon his superior by more fully utilizing his skills and abilities, thereby achieving greater satisfaction from his work through his own efforts. While his efforts in this direction are clearly self-assertive they are genuinely constructive and healthy. These self-assertive activities are quite different in character from those involved in seeking security which are reactive and founded on fear.

Leadership Conditions that Provide for Security. The conditions that lead to a feeling of security on the part of subordinates are summarized by McGregor under three major conditions, namely, a social atmosphere of approval, knowledge, and consistent discipline. These three can be considered as necessary but not sufficient conditions for security. If the individual is threatened from outside the working situation, as by financial or social pressures, or if he is an unstable person emotionally, he may not feel secure in the work situation even though these three conditions are satisfied. Obviously an individual who in his role as subordinate is insecure in relation to his superior cannot, in turn, in his role as superior provide conditions of security for his subordinates. Thus insecure foremen cannot provide the necessary conditions of security for line workers, nor can insecure vice presidents for the department heads under their jurisdiction.

The first condition for security, an atmosphere of approval, is revealed by the general attitude the superior displays toward his subordinates. It is not so much what he does as how he does it. Workers accept orders with less complaint from a foreman whom they feel has their best interest at heart than from one whom they fear and distrust. In order to feel secure the individual must believe that he has the genuine approval of his superior. If the social atmosphere is one of disapproval, or even if it is equivocal, the subordinate can have no assurance that his needs will be satisfied. When the social atmosphere is unfavorable even in-

nocuous actions on the part of the superior will be regarded with suspicion.

Knowledge is an important condition for security because it reduces dependence upon the unpredictable. There are several different kinds of knowledge that are necessary. First of all the subordinate must know the procedures, rules, and regulations that are connected with his job. Without this information he can only learn by trial and error, and the threat of punishment for innocent infractions always hangs over him. Similarly, he needs to have knowledge of his duties and responsibilities, otherwise he does not know when to make a decision. The lack of such knowledge results in buck-passing. Thirdly, the subordinate needs to know something about the over-all policy of the organization. Frequently situations arise that are not covered by formal statements of rules and responsibilities. In such instances unless the subordinate has some general orientation with respect to organizational goals, he will be unable to function adequately. Fourthly, the subordinate needs advance knowledge of changes that affect him or his position. Inertia is greatly enhanced when the worker must forever be prepared for changes. Resistance to change on the part of workers is a common complaint by management. It is not change itself, however, that is resisted but fear of the unknown. Fifthly, the subordinate needs to know how his superior evaluates his performance. Without this information the subordinate may develop a false sense of security. Since his employment is continued and his superior makes no adverse comments concerning his performance, he can only assume that it is satisfactory. Finally, the subordinate needs to know what the personal peculiarities of his superior are. Over and above the formally prescribed duties and procedures for a job, the superior imposes others that reflect his own individual biases and points of view. Some supervisors require originality while others demand strict adherence to rules. Subordinates acquire this knowledge only from long contact with their superiors or by word of mouth from their fellows.

Consistent discipline is the third condition for security. Effective learning requires that the individual know when he is performing correctly and when incorrectly. It is the function of discipline to furnish this information. If discipline is administered only infrequently or incorrectly, the individual will encounter difficulty in understanding what is wanted of him and will feel insecure. While it is generally recognized that discipline involves punishment for incorrect actions, less frequently is its function seen in connection with positive support for correct actions. To be consistent, therefore, discipline must involve an invariable indication of both desirable and undesirable behavior.

Leadership Conditions that Provide for Independence. Participation is the first condition for independence. When the subordinate has the

opportunity to contribute his suggestions and to express his ideas concerning matters that affect or bear upon his work, he is asserting himself as an individual and is showing that he is not wholly dependent upon his superior. Through participation he grows and develops as a person and derives satisfaction from knowing that his ideas are given consideration. Furthermore, by the very process of participation he becomes more and more aware of the problems that confront his superior. He therefore is likely to make fewer demands upon him and to be more understanding of actions that he formerly considered arbitrary.

Responsibility is a second condition of independence. It is a further manifestation on the part of the subordinate of the need to relieve himself of complete dependence upon his superior. Obviously insecure individuals do not seek responsibility. Hence the willingness to assume responsibility is an index of emotional adjustment in the work situation. Too much responsibility, and responsibility suddenly thrust upon the individual, brings about new pressures and insecurities. Therefore the subordinate should not be delegated too much responsibility and it should be delegated to him gradually. It is apparent that there are wide individual differences in capacity to accept responsibility. There are those individuals who are satisfied merely with security and have little need for independence. Others thrive on the inherent risks accompanying responsibility. However, the capacity of subordinates to carry responsibility probably is considerably underestimated. Superiors who themselves are insecure are unlikely to delegate responsibility. Not only are they afraid to run the risks of mistakes on the part of their subordinates, but also they do not wish to permit their subordinates to develop independence to a point where they will be competitors.

The final condition for independence is the right of appeal. There will be instances when superiors and subordinates differ on important questions, and when truly collaborative efforts on the part of both will fail to resolve the disagreement. If the superior gives in to the subordinate he will be following a policy of appeasement that in the long run will cause his subordinate to lose respect for him. On the other hand, if he rejects the subordinate's point of view the subordinate will develop feelings of dependence, and the gains obtained through participation and delegation of responsibility will be undermined. McGregor suggests that the way out of this dilemma is a mechanism of appeal whereby the problem is put for resolution to another individual such as a higher level superior, or to another group such as a grievance committee. By this means the subordinate retains his independence and the superior his integrity. It appears likely that if the subordinate is fully satisfied with respect to his security, participates in matters that are important to him, and has been delegated responsibility, there will be few occasions when

the right of appeal to an independent agent will be necessary. The feelings of security and independence that he has already developed will permit him sufficient personal dignity and understanding so that disagreements either will not arise or will be regarded as quite unimportant.

REFERENCES

1. Benne, K. D., and P. Sheats: Functional roles of group membership, *J. Social Issues,* **4,** No. 2, 41–49, 1948.
2. Blum, M. L.: "Industrial Psychology and Its Social Foundations," Harper, 1949.
3. Browne, C. G.: Study of executive leadership in business. IV. Sociometric pattern, *J. Appl. Psychol.,* **35,** 36–37, 1951.
4. Brunswik, E.: Distal focussing of perception: size-constancy in a representative sample of situations, *Psychol. Monograph,* **56,** No. 254, 1944.
5. Carter, L. F.: An investigation of the relationship between four criteria of leadership ability for three different tasks, *J. Psychol.,* **27,** 245–261, 1949.
6. Carter, L., and M. Nixon: Ability, perceptual, personality, and interest factors associated with different criteria of leadership, *J. Psychol.,* **27,** 377–388, 1949.
7. Carter, L. F., W. Haythorn, and M. Howell: A further investigation of criteria of leadership, *J. Abnormal and Social Psychol.,* **45,** 350–358, 1950.
8. Collins, O., M. Dalton, and D. Roy: Restriction of output and social cleavage in industry, *Appl. Anthrop.,* **5,** No. 3, 1–14, 1946.
9. Cowley, W. H.: Three distinctions in the study of leaders, *J. Abnormal and Social Psychol.,* **23,** 144–157, 1928.
10. Haire, M., and W. F. Grunes: Perceptual defenses: processes protecting an organized perception of another personality, *Human Relations,* **3,** 403–412, 1950.
11. Hemphill, J. K.: "Situational factors in leadership," Ohio State Studies in Personnel, Leadership Studies No. 4, 1949.
12. Hemphill, J. K., and C. M. Westie: The measurement of group dimensions, *J. Psychol.,* **29,** 325–342, 1950.
13. Jacobs, J. H.: The application of sociometry to industry, *Sociometry,* **8,** 181–198, 1945.
14. Jenkins, J. G.: Nominating technique as a method of evaluating air group morale, *J. Aviation Med.,* **19,** 12–19, 1948.
15. Jenkins, W. O.: A review of leadership studies with particular reference to military problems, *Psychol. Bull.,* **44,** 54–79, 1947.
16. Katz, D., N. Maccoby, and N. C. Morse: "Productivity, Supervision, and Morale," University of Michigan, Survey Research Center, 1950.
17. Krech, D., and R. S. Crutchfield: "Theory and Problems of Social Psychology," McGraw-Hill, 1948.
18. Libo, L. M.: Attitude prediction in labor relations—a test of "understanding," Stanford University, Graduate School of Business, Division of Industrial Relations, *Studies in Industrial Relations,* **10,** 1948.
19. Lindley, G. E. (ed.): "Handbook of Social Psychology," Addison-Wesley, 1953.
20. Matthews, J.: "Research in the Development of Valid Situational Tests. I. Survey of the literature," American Institute for Research, 1951.

21. McGregor, D.: Conditions of effective leadership in an industrial organization, *J. Consult. Psychol.*, **8**, 55–63, 1944.
22. Miller, J. (ed.): "Experiments in Social Progress," McGraw-Hill, 1950.
23. Morris, R. T., and M. Seeman: The problem of leadership: an interdisciplinary approach, *Am. J. Sociol.*, **56**, 149–155, 1950.
24. Newcomb, T. M.: "Social Psychology," Dryden, 1950.
25. Roethlisberger, F. J., and W. J. Dickson: "Management and the Worker," Harvard University Press, 1947.
26. Roff, M.: A study of combat leadership in the Air Force by means of a rating scale: group differences, *J. Psychol.*, **30**, 229–239, 1950.
27. Sargent, S. S.: Conceptions of role and ego in contemporary psychology, in Rohrer, J. H., and M. Sherif: "Social Psychology at the Crossroads," Harper, 1951.
28. Shartle, C. L.: Leadership and executive performance, *Personnel*, **25**, No. 5, 370–380, 1949.
29. Smith, M.: Social situation, social behavior, social group, *Psychol. Rev.*, **52**, 224–229, 1945.
30. Stogdill, R. M.: Personal factors associated with leadership: a survey of the literature, *J. Psychol.*, **25**, 35–71, 1948.
31. Stogdill, R. M.: Leadership, membership and organization, *Psychol. Bull.*, **47**, 1–14, 1950.
32. Stogdill, R. M., and C. L. Shartle: Methods for determining patterns of leadership behavior in relation to organizational structure and objectives, *J. Appl. Psychol.*, **32**, 286–291, 1948.
33. Survey Research Center: "Productivity, Supervision, and Employee Morale," University of Michigan, 1948.
34. Wherry, R. J., and D. H. Freyer: Buddy ratings: popularity contest or leadership criteria?, *Personnel Psychol.*, **2**, 147–159, 1949.

Name Index

Subject Index